Niveau(x) d'enseignement : 4 Numéro EDUQ : 9170489
Centre d'intérêt : 755

Québec (Province). Direction générale de l'évaluation et des ressources didactiques. Direction du développement de l'évaluation.

Impact du choix de la langue utilisée dans les épreuves en anglais langue seconde : rapport / Direction du développement de l'évaluation, Direction générale de l'évaluation et des ressources didactiques. -- [Québec] : Ministère de l'éducation, c1990.
20, [20] p. ; 28 cm.

Code : 9091-0345 , 16-7291

ISBN 2-550-15313-8

Descripteur(s) : Anglais (Langue) * Langue seconde * Examen scolaire * Elaboration d'examens * Secondaire (Ordre d'enseignement)

Ce rapport décrit les résultats d'une étude qui visait à mesurer l'impact du choix de la langue utilisée dans les épreuves en anglais langue seconde. La recherche a été réalisée à partir d'une épreuve en anglais langue seconde pour la 2e année du secondaire. Deux versions de cette épreuve ont été administrées à un échantillon de 803 élèves. La première version, dans laquelle les stimuli et les choix de réponse étaient en anglais, a été administrée à 345 élèves. La deuxième version, formulée en français, a été administrée à 458 élèves. Les résultats, traités selon l'approche des biais, ont permis d'établir que la langue utilisée dans les épreuves n'a pas influencé de façon importante la performance des élèves. Les différences observées, quoique non importantes, ont généralement favorisé l'épreuve où la seule langue utilisée était l'anglais. Dans sa conclusion, le rapport recommande que, dans les épreuves d'anglais langue seconde, tout le contenu soit rédigé en anglais.

Type de document :

9776

RAPPORT

Impact du choix de la langue utilisée dans les épreuves en anglais langue seconde

Québec

RAPPORT

Impact du choix de la langue utilisée dans les épreuves en anglais langue seconde

Direction du développement de l'évaluation
Direction générale de l'évaluation et des ressources didactiques

© Gouvernement du Québec
Ministère de l'Éducation, 1990 — 9091-0345

ISBN 2-550-15313-8

Dépôt légal – deuxième trimestre 1990
Bibliothèque nationale du Québec

TABLE DES MATIÈRES

Page

Avant-propos 1

1. La problématique 2

2. La présentation des épreuves 3

3. L'échantillon de la population cible 5

4. Le modèle d'analyse 8

5. La présentation des résultats 11

6. La conclusion et les recommandations 15

Annexes

Avant-propos

Le présent rapport décrit les résultats d'une étude réalisée par une équipe de la Banque d'instruments de mesure (BIM) de la Société GRICS à la demande du ministère de l'Éducation du Québec. Cette étude visait à mesurer l'impact du choix de la langue utilisée dans les épreuves en anglais langue seconde.

La recherche a été réalisée à partir d'une épreuve en anglais langue seconde pour la 2e année du secondaire. Deux versions de cette épreuve ont été administrées à un échantillon d'élèves provenant des différentes régions administratives du Québec. Dans la première version, les stimuli et les choix de réponse étaient en anglais; dans la seconde, ces éléments étaient formulés en français.

Les résultats ont été traités par le groupe BIM selon l'approche de l'étude des biais, procédé statistique éprouvé et reconnu pour les analyses de cette nature. Il s'agissait de vérifier si des élèves ayant un même niveau d'habileté avaient la même chance de réussir un item indépendamment de la langue dans laquelle étaient rédigés les stimuli et les choix de réponse.

Les analyses statistiques permettent d'abord d'établir que les épreuves étaient valables, au sens où on peut se fier aux résultats obtenus. Par la suite, l'étude démontre que la **langue utilisée dans les épreuves n'a pas influencé de façon importante la performance des élèves**. Précisons que les différences observées, quoique non importantes, ont généralement favorisé l'épreuve où la seule langue utilisée était l'anglais.

Dans sa conclusion, le rapport recommande que, dans les épreuves d'anglais langue seconde, tout le contenu des stimuli et des choix de réponse soit rédigé en **anglais**.

Normand Dufour
Maria Carrera
Michel Fréchette

1. La problématique

L'étude qui fait l'objet du présent rapport visait à répondre à certaines questions concernant la langue utilisée dans les épreuves d'anglais langue seconde au secondaire. Dans un contexte d'évaluation, l'utilisation exclusive de la langue seconde (l'anglais) influence-t-elle négativement ou positivement la performance de l'élève à cette épreuve? Dans une situation de compréhension de l'oral ou de l'écrit, la difficulté d'une question est-elle modifiée par le choix de la langue dans laquelle elle est rédigée? L'élève peut-il manifester également, dans une situation comme dans l'autre, son niveau de maîtrise des opérations mentales visées par le programme?

Il faut préciser que les épreuves d'anglais langue seconde comprennent plusieurs parties: d'abord, une mise en situation dans laquelle on précise le contexte de la situation de communication sur laquelle s'appuie l'activité évaluative. On retrouve ensuite un texte, à lire ou à écouter, à partir duquel l'élève devra réaliser des tâches. Viennent enfin les stimuli eux-mêmes et les choix de réponse parmi lesquels l'élève doit identifier la bonne réponse. La controverse porte sur ces deux derniers éléments.

Quand les personnes concernées, dans le milieu scolaire, discutent de l'évaluation de l'anglais langue seconde, leurs positions s'inscrivent dans deux tendances distinctes, voire divergentes, en ce qui touche le choix de la langue dans laquelle les stimuli et les choix de réponse doivent être rédigés. Toutes s'entendent sur le fait que les textes à lire et ceux de la bande sonore doivent être en anglais.

Une première tendance stipule que les items d'une épreuve en anglais langue seconde doivent être rédigés dans la langue maternelle des candidats, en l'occurrence le français. Par ce moyen, on veut éviter d'influencer négativement la performance de l'élève dans le cas où ce dernier n'aurait pas bien compris les stimuli et les choix de réponse.

Selon la deuxième tendance, ces deux derniers éléments de l'épreuve doivent être rédigés dans la langue seconde, en l'occurrence l'anglais. Ainsi, la situation d'évaluation se déroule entièrement dans la langue seconde, reflétant plus fidèlement les orientations du programme.

Consciente de la difficulté de concilier ces deux positions et de la nécessité d'appuyer son choix sur des informations aussi objectives que possible, la Direction du développement de l'évaluation du MEQ a proposé au milieu scolaire de mener une étude sur cette question. Devant la réaction positive des personnes soutenant l'une ou l'autre tendance, le MEQ a confié au groupe BIM le mandat de réaliser cette recherche à partir de l'épreuve d'appoint de la 2e année du secondaire.

2. La présentation des épreuves

Afin de mener à bien cette étude, le groupe BIM a été mandaté pour produire une épreuve avec la collaboration de quelques commissions scolaires.

Il convient de souligner la participation de madame Doreen Hughes ainsi que celle des enseignants et enseignantes des commissions scolaires suivantes: Victoriaville, Warwick et Jean Rivard. Pour leur apport significatif sur le plan de la révision des épreuves, nous tenons également à remercier les conseillers pédagogiques des commissions scolaires Chambly, Trois-Rivières, Des Cantons, Beauport et Saint-Jean-sur-Richelieu. Soulignons enfin la contribution de monsieur Jonathan Munro Jones, responsable du programme concerné à la Direction de la formation générale. Pour le groupe BIM, c'est madame Diane Provençal qui a supervisé la production des épreuves et qui a effectué les révisions docimologiques appropriées.

L'épreuve d'appoint utilisée pour l'étude permet de vérifier la **maîtrise des opérations mentales et des notions** prescrites par le programme d'anglais langue seconde de deuxième secondaire. Elle est fondée sur la définition de domaine du MEQ portant sur le programme d'anglais 134-214.

Les tableaux-synthèse des deux volets «compréhension», soit celle de l'oral et celle de l'écrit, sont présentés à l'**annexe 1 (1A et 1B)**. Mentionnons que l'épreuve d'appoint produite dans le cadre de cette recherche comportait quatre volets de façon à couvrir complètement le programme. Par contre, seuls les résultats des volets de compréhension ont été considérés, les volets de «production» n'étant pas susceptibles d'être influencés par les résultats de cette recherche. L'épreuve est actuellement disponible par l'entremise de BIM; on peut également en demander une copie au MEQ.

L'épreuve a été bâtie autour d'un **scénario** qui assurait une relation thématique d'un volet à l'autre et qui rendait la situation de communication plausible pour l'élève. Ce faisant, une attention particulière a été accordée à l'indépendance locale des items. En effet, il importait de vérifier si la performance des élèves à l'un ou à l'autre des «volets expérimentaux» n'était pas influencée par le fait d'avoir passé un autre volet.

Afin de bien cerner la problématique, il a été convenu d'administrer deux versions de cette épreuve à deux échantillons représentatifs de la population cible. Il importe de préciser que, dans les deux versions de l'épreuve, tout le contenu est exactement le même: les consignes et les mises en situation sont en français; les textes (à écouter ou à lire) sont en anglais. La distinction se situe uniquement sur le plan de la langue utilisée dans les stimuli et dans les choix de réponse. Le tableau ci-dessous présente la différence entre les deux versions de l'épreuve.

Tableau 1

Organisation des deux versions de l'épreuve

	Consignes et Mises en situation	Textes	Questions	Choix de réponse
Version 1	Français	Anglais	**Anglais**	**Anglais**
Version 2	Français	Anglais	**Français**	**Français**

Le type d'item qui a été privilégié dans l'épreuve est celui des questions formulées avec des choix de réponse. Ce procédé facilite l'administration et, surtout, assure une fidélité à la correction. Les 20 items considérés pour cette étude se répartissent ainsi:

- Compréhension de l'oral : items 1 à 11.
- Compréhension de l'écrit : items 12 à 20.

Les deux versions de l'épreuve ont été administrées le 15 juin 1989. La durée prévue pour l'administration des deux volets en compréhension était d'environ une heure. L'usage du dictionnaire et de toute autre documentation était interdit.

3. L'échantillon de la population cible

L'échantillon a été constitué par le ministère de l'Éducation du Québec selon l'approche utilisée en évaluation de programme. Un total de 100 écoles secondaires privées et publiques des différentes régions administratives du Québec ont été sélectionnées de façon aléatoire. Le nom d'un élève par école a été choisi au hasard. L'épreuve devait être administrée au groupe-classe auquel appartenait cet élève. Ainsi, l'échantillon sélectionné par le MEQ devait être constitué d'environ 1 000 élèves pour chacune des versions. Le tableau suivant présente le résultat de la constitution de l'échantillon.

Tableau 2

Écoles secondaires publiques et privées, sélectionnées par le MEQ

Région administrative	Version 1 anglaise		Version 2 française		Total
	École publique	École privée	École publique	École privée	
01 Bas St-Laurent-Gaspésie	4		3	1	8
02 Saguenay - Lac-St-Jean	2		2	1	5
03 Québec	5	4	7	1	17
04 Trois-Rivières	4		2	3	9
05 Estrie	2	1	2		5
06 Montréal, Montérégie, Laval-Laurentides	16	6	18	3	43
07 Outaouais	1		2		3
08 Abitibi - Témiscamingue	2	1	2		5
09 Côte-Nord	1	1	3		5
Total	37	13	41	9	100
	50		50		

Parmi les cent écoles sélectionnées, 34 ont accepté de participer à l'expérimentation. Le tableau suivant montre la répartition des écoles participantes dans les différentes régions administratives du Québec.

Tableau 3

Nombre d'écoles secondaires publiques et privées ayant participé à l'expérimentation

Région administrative	Version 1 anglaise	Version 2 française
01 Bas St-Laurent - Gaspésie		3
02 Saguenay - Lac-St-Jean		3
03 Québec	4	3
04 Trois-Rivières	2	3
05 Estrie	1	1
06 Montréal, Montérégie, Laval-Laurentides	4	4
07 Outaouais	2	1
08 Abitibi-Témiscamingue	1	1
09 Côte-Nord	1	
TOTAL	15	19

À première vue, le nombre de participants peut paraître peu élevé par rapport au nombre d'écoles sélectionnées; plusieurs raisons permettent d'expliquer cet écart:

- Toute l'information et la documentation concernant cette épreuve d'appoint ont été transmises par le groupe BIM. Même s'il était précisé que l'opération était réalisée à la demande du MEQ et que les personnes rejointes faisaient partie de l'échantillon établi par le MEQ, les responsables dans les commissions scolaires n'ont pas senti la même obligation de participer que si la demande était venue directement du MEQ.

- Comme cela a été précisé au début de ce rapport, le milieu scolaire est divisé par rapport à l'objet même de cette recherche. Aussi, la procédure d'échantillonnage a fait en sorte que des tenants d'une des tendances ont été invités à administrer une épreuve associée à l'autre tendance, ce qui a pu accroître le nombre d'abstentions.

- L'épreuve finale a été prête plus tard que prévu. Certains organismes avaient déjà mis en place une autre procédure d'évaluation.

- Le hasard a fait que plusieurs élèves ont été sélectionnés dans des «classes spéciales», si bien que les enseignants concernés nous ont informés qu'il leur était impossible d'utiliser cette épreuve parce qu'elle dépassait les capacités réelles de leurs élèves.

- Quelques organismes ont administré les épreuves et analysé les résultats pour leurs besoins, mais ils ont détruit les données de base avant de les transmettre au groupe BIM. Nous estimons qu'une centaine de résultats ont ainsi été perdus.

Le tableau suivant présente le nombre d'élèves qui ont participé à l'expérimentation.

Tableau 4

Nombre d'élèves ayant participé à l'expérimentation

Région administrative	Version 1 anglaise	Version 2 française	Total
01 Bas St-Laurent - Gaspésie		47	47
02 Saguenay - Lac-St-Jean		87	87
03 Québec	73	108	181
04 Trois-Rivières	28	74	102
05 Estrie	34	30	64
06 Montréal, Montérégie, Laval-Laurentides	111	53	164
07 Outaouais	56	32	88
08 Abitibi - Témiscamingue	21	27	48
09 Côte-Nord	22		22
TOTAL	345	458	803

Les responsables des 34 écoles participantes ont expédié les cahiers de réponses des élèves au groupe BIM, qui a alors procédé à la compilation et au traitement statistique des résultats. Le tableau 4 montre que l'échantillon était constitué de 803 élèves: 345 pour la version 1, entièrement en anglais, et 458 élèves pour la version 2, dont une partie était en français. On remarque également que les participants sont répartis différemment d'une région à l'autre; en effet, on peut regretter de n'avoir eu aucune réponse des régions 01 et 02; par ailleurs, les participants pour la version française de la région 06 auraient pu être plus nombreux. Quoi qu'il en soit, l'échantillon était tout de même suffisant pour assurer la validité des résultats.

4. Le modèle d'analyse

Afin de répondre adéquatement aux questions soulevées dans cette recherche, il était nécessaire de procéder à deux types d'analyse complémentaires. Il fallait d'abord établir la valeur des épreuves, c'est-à-dire examiner le niveau de difficulté des items, vérifier si les résultats obtenus étaient plausibles et, le cas échéant, éliminer les items pour lesquels beaucoup de résultats incohérents auraient été observés. Ensuite, il fallait mesurer l'écart entre la performance des élèves ayant un même niveau d'habileté, pour chacune des deux versions de l'épreuve.

Étant donné la disparité de la répartition des sujets dans les deux échantillons, le groupe BIM a privilégié des méthodes d'analyse inspirées de la théorie des réponses aux items (TRI)[1]. Cette approche est tout à fait appropriée notamment parce que les paramètres des items sont estimés, d'une façon indépendante des sujets et donc de leur niveau d'habileté. Cependant, pour faciliter une compréhension globale de la situation, on présente à l'annexe 3 les statistiques classiques concernant les résultats obtenus dans les deux échantillons; on y constate que les résultats sont très semblables. À l'annexe 3B, on voit jusqu'à quel point la proportion de choix des leurres est similaire d'une version à l'autre.

---

[1]C.L. HULIN, F. DRASGOW ET C.K. PARSONS., Item Response Theory: Applications to Psychological Measurement, Homewood, Ill. Dow Jones-Irwin, 1983.

Dans un premier temps, les résultats ont été traités à l'aide du programme MICROSCALE. Ce programme, où l'on utilise un seul paramètre (celui de la **difficulté des items**), permet de vérifier ce qu'on pourrait appeler le «degré d'homogénéité» des réponses des élèves. Si un patron de réponses est incohérent (par exemple, un élève fort qui échoue à plusieurs items faciles), on peut repérer l'item qui fait problème et l'éliminer au besoin.

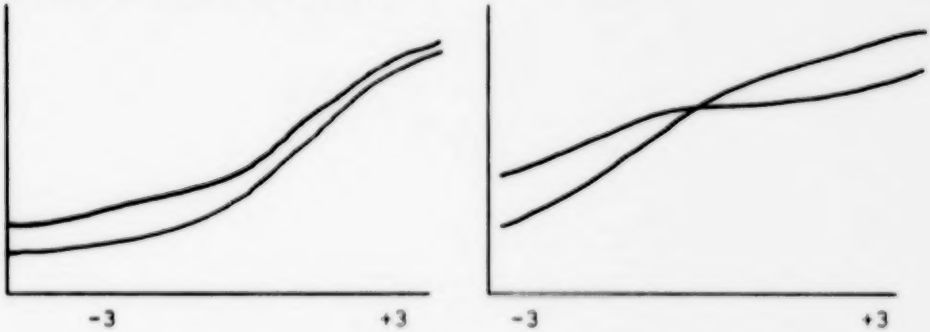
Nous avons ensuite déterminé l'indice de difficulté de chaque item. Ces indices sont exprimés en valeurs INDIBIM. Ils varient de 140 à 260, allant du plus facile au plus difficile. L'annexe 2 présente les indices de difficulté pour chacun des items. Il est intéressant de noter que la moyenne des indices de difficulté est identique pour chaque version de l'épreuve.

Aucun item n'a été éliminé car les patrons de réponse étaient cohérents. Ces analyses statistiques ont démontré que l'**épreuve était fiable** et que les indices calculés étaient valables sur le plan statistique.

Dans un deuxième temps, il s'agissait de déterminer pour chaque item l'importance des écarts entre les niveaux de performance des sujets d'habileté ou de compétence semblable. Pour cela, nous avons recouru à un modèle logistique à trois paramètres, soit la difficulté de l'item, la discrimination et l'effet dû au hasard. Le logiciel Bilog a été utilisé à cet effet. Cette analyse est fondée sur la comparaison de la **courbe caractéristique d'un item** (CCI), celle-ci étant estimée successivement à l'aide de deux ou plusieurs échantillons. Ainsi, pour chacun des vingt items, on retrouve deux courbes décrivant la performance des élèves pour chacune des versions de l'épreuve.

Cette méthode est très intéressante pour plusieurs raisons. Premièrement, elle fournit des indices qui permettent de vérifier si l'écart entre la performance des élèves pour chacune des versions est important ou non. Deuxièmement, l'analyse des deux courbes caractéristiques permet au lecteur profane en statistiques d'observer visuellement l'importance de l'écart.

Pour bien comprendre cette approche, utilisons un exemple. Sur les graphes suivants, la courbe indique en ordonnée la probabilité qu'a un individu ayant une habileté donnée (présentée en abscisse) de réussir un item. Plus on se déplace vers la droite, plus l'individu a un niveau d'habileté supérieur. Plus la courbe est haute, plus un individu a de chance de réussir un item.



Supposons deux individus ayant un même niveau d'habileté; si celui à qui on administre l'épreuve en anglais a une plus grande chance de réussir un item, on dira que cet item est entaché d'un biais en faveur de la version anglaise. Ainsi, la note obtenue par un élève dépendrait non pas seulement de sa compétence (en regard de l'objet mesuré) mais **aussi de la langue utilisée** dans l'épreuve.

Comme la courbe caractéristique d'un item permet d'estimer cette probabilité pour chacun des niveaux d'habileté (variant théoriquement de - l'infini à + l'infini, mais généralement compris entre -3 et +3), on se trouve à mesurer cet écart (biais) pour l'ensemble d'un groupe. D'ailleurs, comme le démontre l'exemple ci-dessus, le biais n'est pas toujours identique pour tous les individus; il peut même y avoir un effet (biais) inverse pour des élèves forts et pour des élèves faibles.

Pour mesurer l'écart entre deux sous-populations, on superpose sur un même graphe les deux courbes de l'item. L'espace entre ces deux courbes correspond à la différence que l'on veut mesurer. On comprendra alors que plus l'aire entre les deux courbes est grande, plus il y a de chance que l'item soit entaché d'un biais. À l'aide des statistiques, on peut quantifier l'aire qui sépare les deux courbes et déterminer si cet écart est important ou non.

5. La présentation des résultats

Les courbes caractéristiques de tous les items de l'épreuve sont présentées à l'**annexe 4**. La courbe caractéristique des items de la version française est représentée par des <u>astérisques</u>. Les <u>signes +++</u> correspondent aux courbes caractéristiques des items rédigés uniquement en anglais.

Globalement, nous n'avons trouvé aucune différence importante entre la performance des deux sous-populations; c'est ce qu'indique le tableau 5 ci-dessous, qui présente les valeurs **RMSE**, indice permettant de mesurer le biais d'un item[7]. Cet indice est basé sur l'écart entre les deux courbes d'un même item. Explicitement, le RMSE est la racine carrée de la moyenne des différences au carré entre les deux courbes pour 601 niveaux d'habileté espacés entre - 3,0 et + 3,0. Comme il est d'usage, les valeurs RMSE ont subi la transformation suivante: ($X\ 10^{-2}$).

<u>Tableau 5</u>

Indice de biais

Compr. de l'oral		Compr. de l'écrit	
Item	RMSE	Item	RMSE
1	0,19	12	0,65
2	0,12	13	0,36
3	0,15	14	0,49
4	0,13	15	0,21
5	0,36	16	0,38
6	0,36	17	0,27
7	0,35	18	0,41
8	0,38	19	0,47
9	0,41	20	0,37
10	0,53		
11	0,58		

[7]Ibid.

Pour bien interpréter ce tableau, il faut savoir que la moyenne des indices RMSE est de 0,36 et l'écart type de 0,15. De plus, la taille de l'échantillon n'a pas permis de déterminer un seuil critique pour les valeurs RMSE. Cependant, on peut analyser les items pour lesquels les valeurs sont vraiment supérieures à un écart type de la moyenne car ils sont plus susceptibles d'être entachés d'un biais. On constate que très peu d'items entrent dans cette catégorie; il s'agit des items 2, 4, 11 et 12. Dans certains cas, notamment pour l'item 2, l'observation des deux courbes à l'annexe 4 ne laisse pas croire à une différence importante.

Nous devons conclure que la **langue utilisée dans les épreuves d'anglais langue seconde n'a pas influencé la performance des élèves**.

Comme on peut le constater en examinant les différents graphes de l'annexe 4, la performance des élèves a tout de même été un peu différente pour quelques items. Même si ces différences ne sont pas très importantes, il apparaît approprié de décrire sommairement les observations faites à partir de ces items.

Compréhension de l'oral

En compréhension de l'oral, on a observé des écarts légers dans le cas des items 5, 6, 7, 8, 9, 10 et 11.

La probabilité de réussir l'**item 5** a été plus élevée dans la version anglaise de l'épreuve. L'item 5 mesurait l'opération mentale de la découverte par rapport à une intention de s'informer sur des intérêts, des attitudes et des sentiments.

Quant à l'**item 6**, on a constaté chez les élèves dont le niveau d'habileté était faible ou moyen, une meilleure probabilité de le réussir dans la version anglaise de l'épreuve. L'item 6 évaluait l'opération mentale de la déduction en vue de s'informer sur des gestes à poser.

En ce qui a trait à l'**item 7**, on a observé un écart particulier entre les deux versions. Chez les élèves dont le niveau d'habileté était très faible, la probabilité de réussir l'item était plus élevée chez ceux qui ont passé la version anglaise de l'épreuve. Cependant, le contraire s'est produit chez les élèves dont le niveau d'habileté était moyen

ou supérieur. Ces individus ont témoigné une meilleure probabilité de réussite avec la version française de l'épreuve. L'item 7 mesurait l'opération mentale de la recherche en vue de s'informer sur des faits relatifs à une personne, à un lieu et à des choses.

Pour ce qui concerne l'**item 8**, l'écart observé n'a pas été constant dans les deux sous-populations. Chez les élèves dont le niveau d'habileté était faible, la probabilité de réussir l'item 8 a été plus élevée avec la version anglaise de l'épreuve. Cependant, les élèves possédant un niveau d'habileté moyen ou supérieur ont mieux réussi avec la version française de l'épreuve. L'item 8 évaluait l'opération mentale de la déduction en vue de s'informer sur des intérêts, des attitudes et des sentiments.

À l'**item 9**, on a observé un écart constant dans les deux sous-populations. La probabilité de réussir a été supérieure chez les élèves qui ont passé la version anglaise de l'épreuve. L'item 9 mesurait l'opération mentale de la découverte en vue de s'informer sur des intérêts, des attitudes et des sentiments.

La probabilité de réussir l'**item 10** a été supérieure chez les élèves qui ont passé la version française de l'épreuve. L'item 10 visait à mesurer l'opération mentale du repérage en vue de s'informer sur des faits relatifs à une personne, un lieu et des choses.

Quant à l'**item 11**, l'écart observé n'a pas été constant dans les deux sous-populations. Chez les individus dont le niveau d'habileté était faible ou moyen, la probabilité de réussir l'item 11 a été supérieure dans la première sous-population, soit chez les élèves qui ont passé la version anglaise de l'épreuve. Cependant, chez les individus dont le niveau d'habileté était supérieur à la moyenne, la probabilité de réussir a été plus élevée dans la deuxième sous-population, soit chez les élèves qui ont passé la version française.

Compréhension de l'écrit

Dans le volet compréhension de l'écrit, on a observé des écarts dans le cas des items 12, 14, 16, 18, 19 et 20.

En ce qui a trait à l'**item 12**, un écart plus considérable a été noté dans les deux sous-populations. La probabilité de

réussir a été supérieure chez les élèves à qui la version anglaise de l'épreuve a été administrée. L'item 12 mesurait l'opération mentale de la déduction en vue de s'informer sur des intérêts, des attitudes et des sentiments.

À l'**item 14**, le même phénomène s'est produit. La probabilité de réussir cet item a été supérieure dans la première sous-population, soit chez les élèves qui ont passé la version anglaise de l'épreuve. L'item 14 évaluait l'opération de la déduction en vue de s'informer sur des intérêts, des attitudes et des sentiments.

L'**item 16** mesurait l'opération mentale de la déduction en vue de s'informer sur des faits relatifs à une personne, un lieu et des choses. La probabilité de réussir l'item 16 a été plus élevée chez les élèves qui ont passé la version anglaise.

En ce qui a trait à l'**item 18**, la probabilité de réussir a été plus élevée chez les élèves qui ont passé la version française de l'épreuve. Ce phénomène a été observé plus particulièrement chez les individus dont le niveau d'habileté était moyen ou supérieur. L'item 18 mesurait l'opération de la déduction en vue de se distraire.

Quant à l'**item 19**, un écart a été observé chez les individus dont le niveau d'habileté était faible ou moyen. Ces derniers ont témoigné une probabilité supérieure de réussir dans la version anglaise de l'épreuve. L'item 19 évaluait l'opération de la comparaison en vue de s'informer sur des intérêts, des attitudes et des sentiments.

L'**item 20** mesurait l'opération de la comparaison en vue de s'informer sur des intérêts, des attitudes et des sentiments. Un phénomène différent a été observé; chez les individus dont le niveau d'habileté était faible ou moyen, la probabilité de réussir a été plus élevée dans la version anglaise de l'épreuve. Cependant, chez les individus dont le niveau d'habileté était supérieur, la probabilité de réussir a été plus élevée dans la version française de l'épreuve.

6. La conclusion et les recommandations

L'étude décrite ici visait à déterminer si l'utilisation de la langue anglaise dans la rédaction des items d'une épreuve en anglais langue seconde exerçait un impact significatif sur la performance des élèves. Pour vérifier cette hypothèse, on a administré deux versions d'une même épreuve à un échantillon représentatif de 952 élèves des différentes régions administratives du Québec. Pour la version 1, les stimuli et les choix de réponse étaient en anglais. Pour la version 2, ces éléments étaient en français.

Les résultats obtenus ont confirmé qu'il n'y avait **pas de différence importante** entre la performance de chacun des deux sous-groupes de la population cible.

Par conséquent, étant donné qu'il n'y a pas de différence, on peut dire que les stimuli et les choix de réponse peuvent tout aussi bien être rédigés en anglais ou en français.

Il faudra considérer d'autres aspects pour décider de privilégier une langue plutôt qu'une autre. Parmi ceux-ci, notons que l'approche communicative préconise l'intégration des fonctions langagières dans un contexte authentique et vraisemblable. Elle vise autant la compréhension que la production d'un message réel, et cela, autant dans les situations d'évaluation que dans les situations d'apprentissage.

Dans ce contexte, il est recommandé que **les stimuli et les choix de réponse des épreuves en anglais langue seconde soient entièrement rédigés dans la langue seconde, en l'occurrence l'anglais**. En plus de ne causer aucun préjudice à l'élève, l'approche communicative véhiculée dans le programme d'études serait respectée plus rigoureusement.

Annexe 1A

TABLEAU-SYNTHESE DES DIMENSIONS

Volets compréhension
2e secondaire

Opérations / Intentions		Repérage	Comparaison	Déduction	Découverte	Recherche	% CO	% CE
S'informer sur des faits relatifs à une personne, un lieu, des choses...		1) #10	2)	3) #1 #11 #16	4)	5) #7	9	
S'informer sur des gestes à poser.		6)	7)	8) #3 #4 #6	9)	10)	8	
S'informer sur des intérêts, des attitudes, des sentiments.		11)	12) #19 #20	13) #2 #8 #12 #13 #14	14) #5 #9	////// ////// ////// //////	8	7
Se distraire.		15)	16)	17) #18	18) #17	19) #15	8	7
%	CO	10	3	8	8	4	33	27
	CE	9	3	6	6	3		

////// : Cette dimension n'existe pas. Elle ne fera donc pas l'objet de mesure.

Les chiffres suivis d'une parenthèse représentent les dix-neuf dimensions de la définition de domaine. Pour faciliter la compréhension, nous avons enlevé ces dimensions dans le tableau de l'annexe 1B.

Annexe 1B

TABLEAU-SYNTHÈSE DES DIMENSIONS

Volets compréhension
2e secondaire

Opérations / Intentions		Repérage	Comparaison	Déduction	Découverte	Recherche	% CO	% CE
S'informer sur des faits relatifs à une personne, un lieu, des choses...		#10		#1 #11 #16		#7		
S'informer sur des gestes à poser.				#3 #4 #6				
S'informer sur des intérêts, des attitudes, des sentiments.			#19 #20	#2 #8 #12 #13 #14	#5 #9	//////// //////// //////// ////////		
Se distraire.				#18	#17	#15		
%	CO							
	CE							

//////// : Cette dimension n'existe pas. Elle ne fera donc pas l'objet de mesure.

Annexe 2

Valeurs INDIBIM pour chaque version.

Numéro de l'item	Version anglaise	Version française
# 1	192	190
# 2	199	200
# 3	192	194
# 4	179	170
# 5	203	208
# 6	190	192
# 7	197	189
# 8	194	189
# 9	208	214
# 10	204	194
# 11	218	224
# 12	201	206
# 13	198	197
# 14	201	207
# 15	206	211
# 16	206	214
# 17	203	200
# 18	207	200
# 19	199	201
# 20	201	201

Annexe 3A

Comparaison des moyennes

	Version anglaise		Version française	
	Moyenne	écart type	Moyenne	écart type
Score total	14,073	4,26	12,44	3,73
Oral	8,047	2,586	7,41	2,28
Écrit	6,026	5,04	2,11	2,117

Annexe 3B

Proportion du choix des leurres

Compréhension de l'oral

	Version anglaise					Version française				
Item	A	B	C	Blanc	Rbis	A	B	C	Blanc	Rbis
1	82,6*	4,3	12,8	0,3	,81	81,7*	2,8	14,4	1,1	,44
2	8,1	17,7	73,9*	0,3	,74	3,1	30,8	65,1*	1,1	,53
3	83,2*	4,6	11,9	0,3	,83	76,0*	8,5	14,4	1,1	,49
4	1,4	93,6*	4,6	0,3	,77		96,1*	2,8	1,1	,63
5	2,3	66,1*	31,3	0,3	,70	2,0	48,9*	48,0	1,1	,46
6	11,9	2,6	85,2*	0,3	,59	14,4	6,8	77,3*	1,5	,45
7	4,1	18,8	76,8*	0,3	,70	7,0	9,8	82,1*	1,1	,56
8	80,0*	12,8	7,0	0,3	,28	82,3*	7,6	9,0	1,1	,40
9	57,1*	4,6	37,7	0,6	,57	36,0*	5,2	57,6	1,1	,46
10	65,2*	21,4	13,0	0,3	,71	75,1*	10,0	13,8	1,1	,54
11	38,6*	49,0	12,2	0,3	,43	20,1*	54,1	24,7	1,1	,17

Annexe 3B (suite)

Compréhension de l'écrit

	Version anglaise							Version française						
Item	A	B	C	D	E	Blanc	Rbis	A	B	C	D	E	Blanc	Rbis
12	20,3	0,6	69,3*	1,2	8,4	0,3	,25	26,2	0,2	51,3*	2,2	9,2	0,9	,17
13	4,1	74,2*	7,8	0,6	13,0	0,3	,44	3,7	70,3	10,5	0,9	13,3	1,3	,29
14	18,0	0,3	3,2	69,9*	8,4	0,3	,55	34,5	0,7	5,5	50,4*	7,9	1,1	,37
15	18,0	60,3*	4,1	17,1		0,6	,67	26,9	41,9*	8,1	21,8		1,3	,46
16	2,3	3,5	32,2	61,7*		0,3	,66	1,7	6,6	54,6	36,0		1,1	,52
17	10,7	5,5	17,4	65,8		0,6	,58	8,7	3,5	22,5	64,2*		1,1	,37
18	20,3	20,0	58,8*			0,9	,27	22,7	11,4	64,4*	0,2		1,3	,29
19	8,4	10,4	8,1	71,9*		1,2	,37	8,5	13,1	15,1	62,4*		0,9	,36
20	69,0*	1,7	22,0	5,8		1,4	,31	62,4*	3,3	26,4	5,8	0,2	0,9	,30

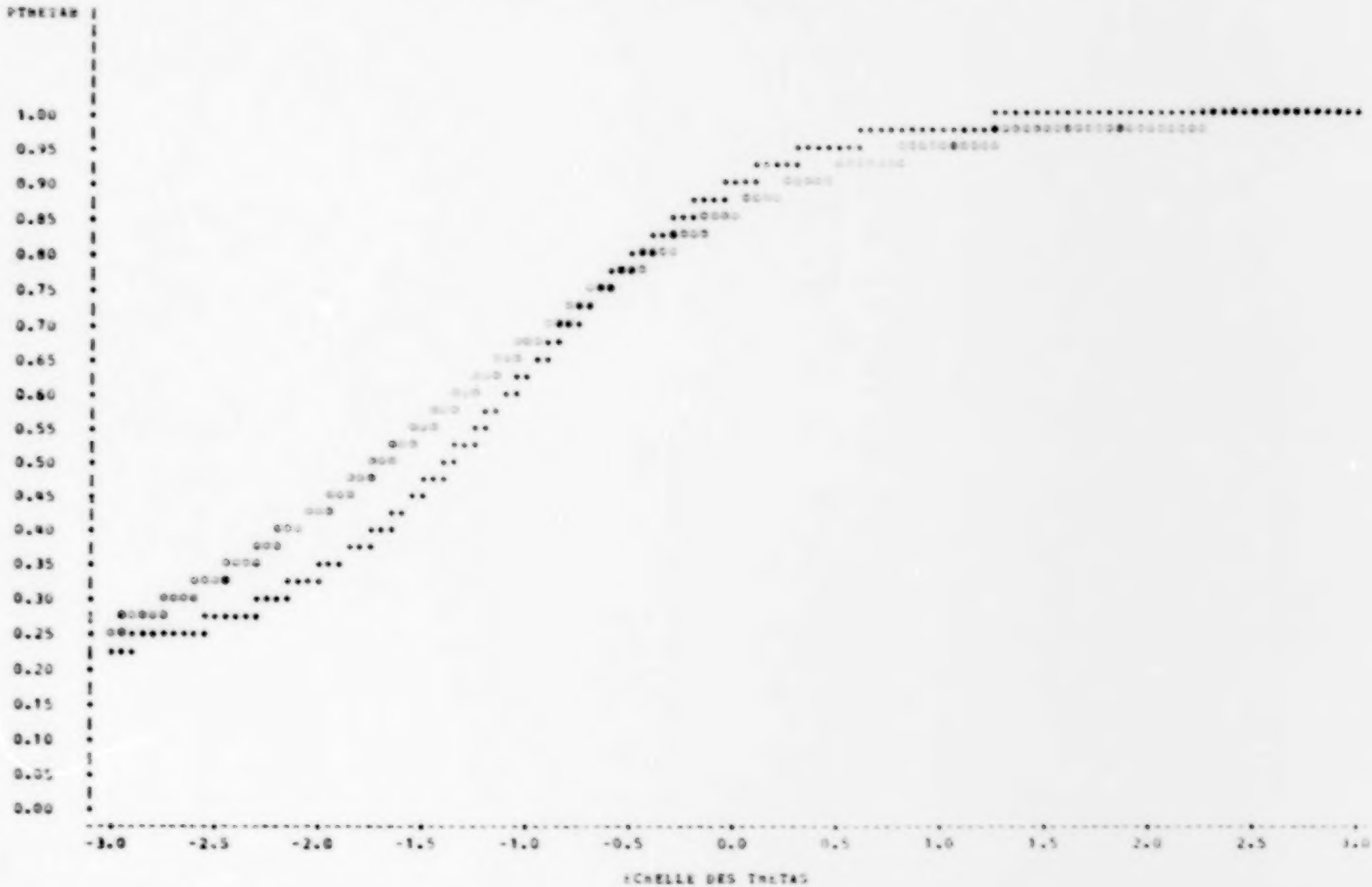
Annexe 4A

RMSE : ANGLAIS VS FRANCAIS   13:55 SATURDAY, NOVEMBER 4, 1989   1

ITEM : 0001   RMSE : 0.00[illegible]

PARAMETRES (BASE) : A= 0.67929 B= -1.41101 C= 0.13913

PARAMETRES COMPARAISON : A= 1.04243 B= -1.07467 C= 0.20764

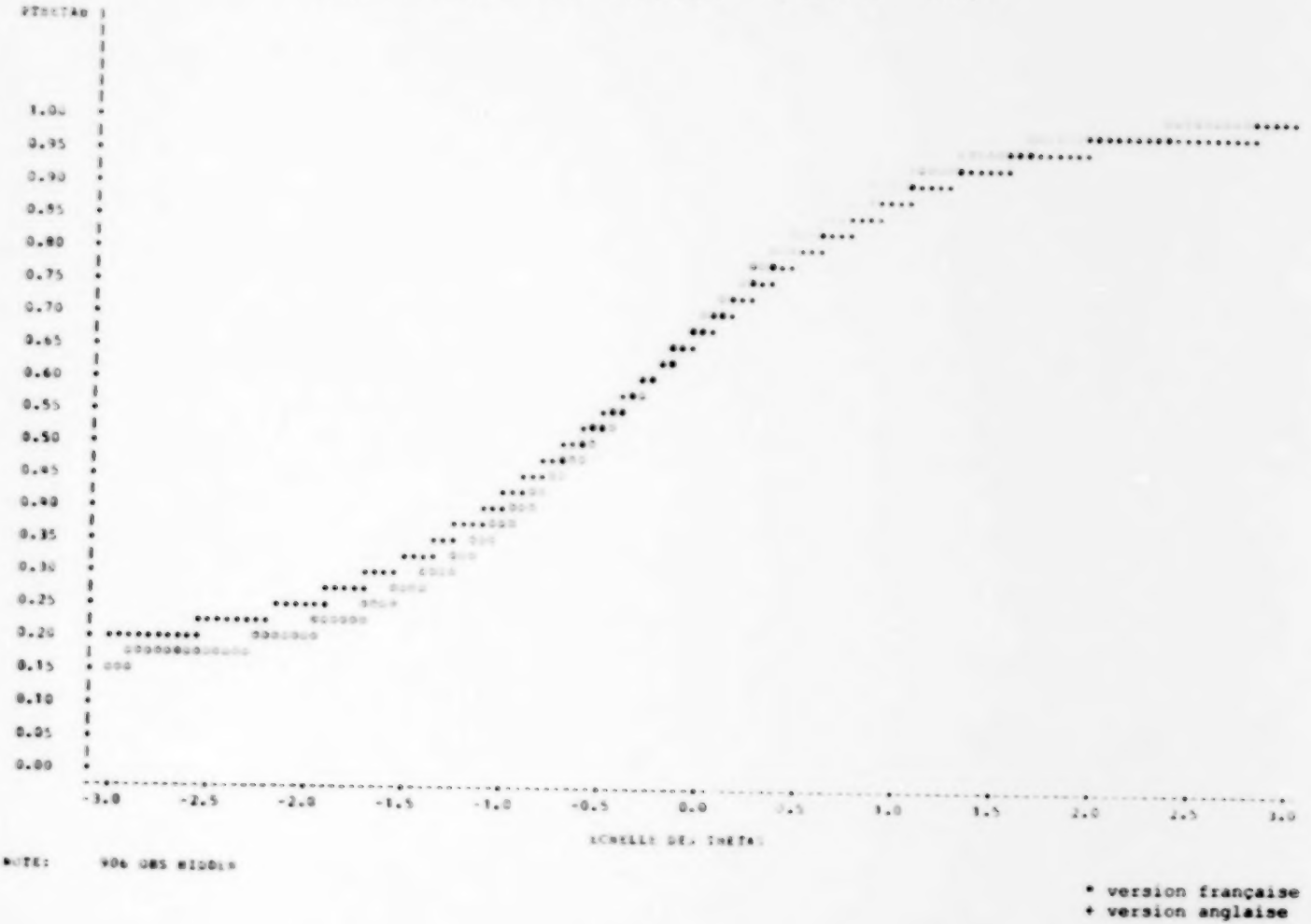


NOTE: 910 OBS HIDDEN

* version française

\+ version anglaise

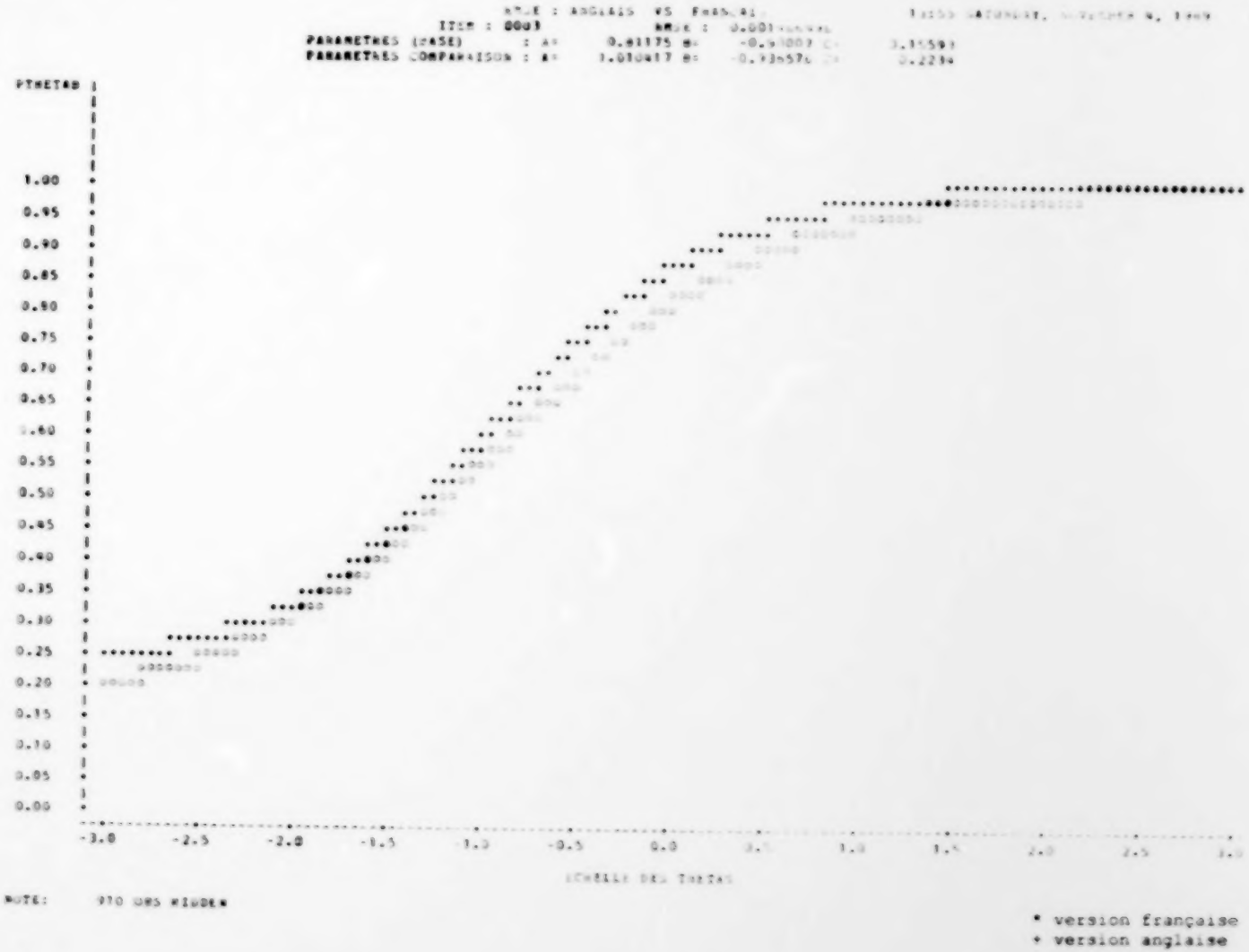
Annexe 4B



NOTE: 906 OBS MISES

* version française
\+ version anglaise

Annexe 4C


ITEM : 0003
PARAMETRES (CASE) : A= 0.831175
PARAMETRES COMPARAISON : A= 1.010417
PTHETAD
1.00
0.95
0.90
0.85
0.80
0.75
0.70
0.65
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
-3.0
-2.5
-2.0
-1.5
-1.0
-0.5
0.0
1.0
1.5
2.0
2.5
3.0
ECHELLE DES THETAS
NOTE: 970 OBS HIDDEN
* version française
+ version anglaise

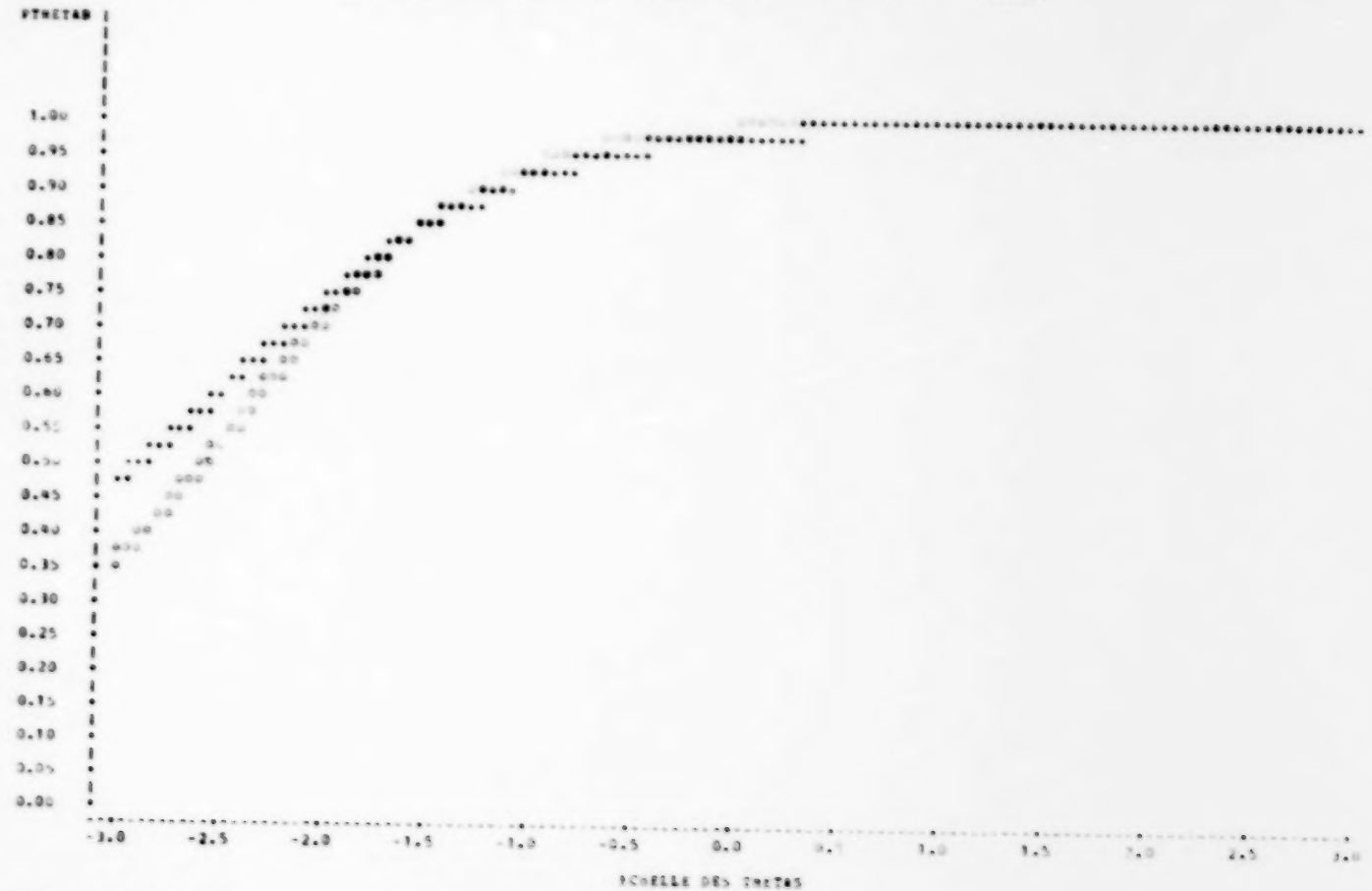
Annexe 4D

AMSE : ANGLAIS VS FRANCAIS

ITEM : 0004 AMSE : 0.0913

PARAMETRES (BASE) : A= 1.04362 B= -2.19184 C= 0.19016

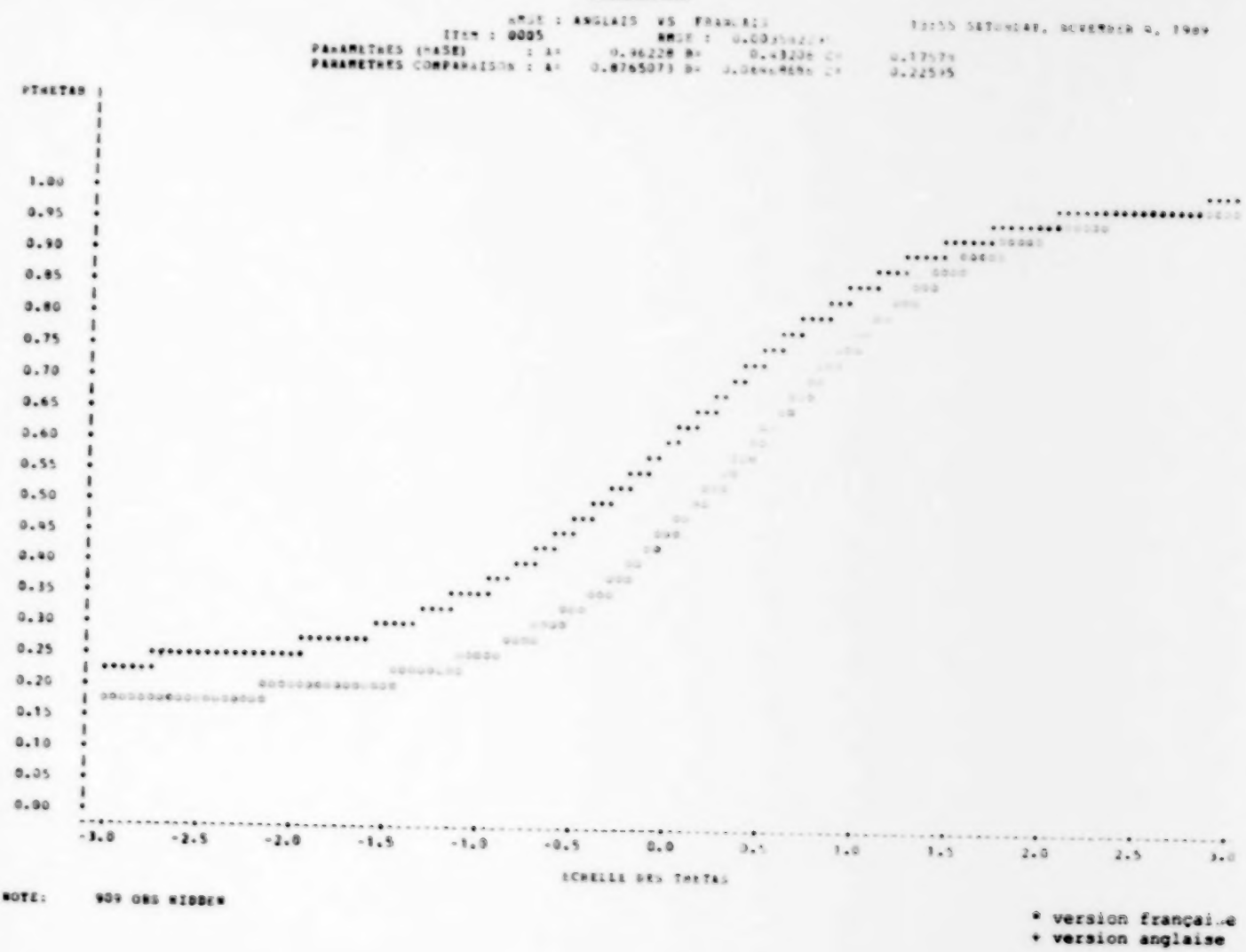
PARAMETRES COMPARAISON : A= 0.8669749 B= -2.4 C= 0.23056



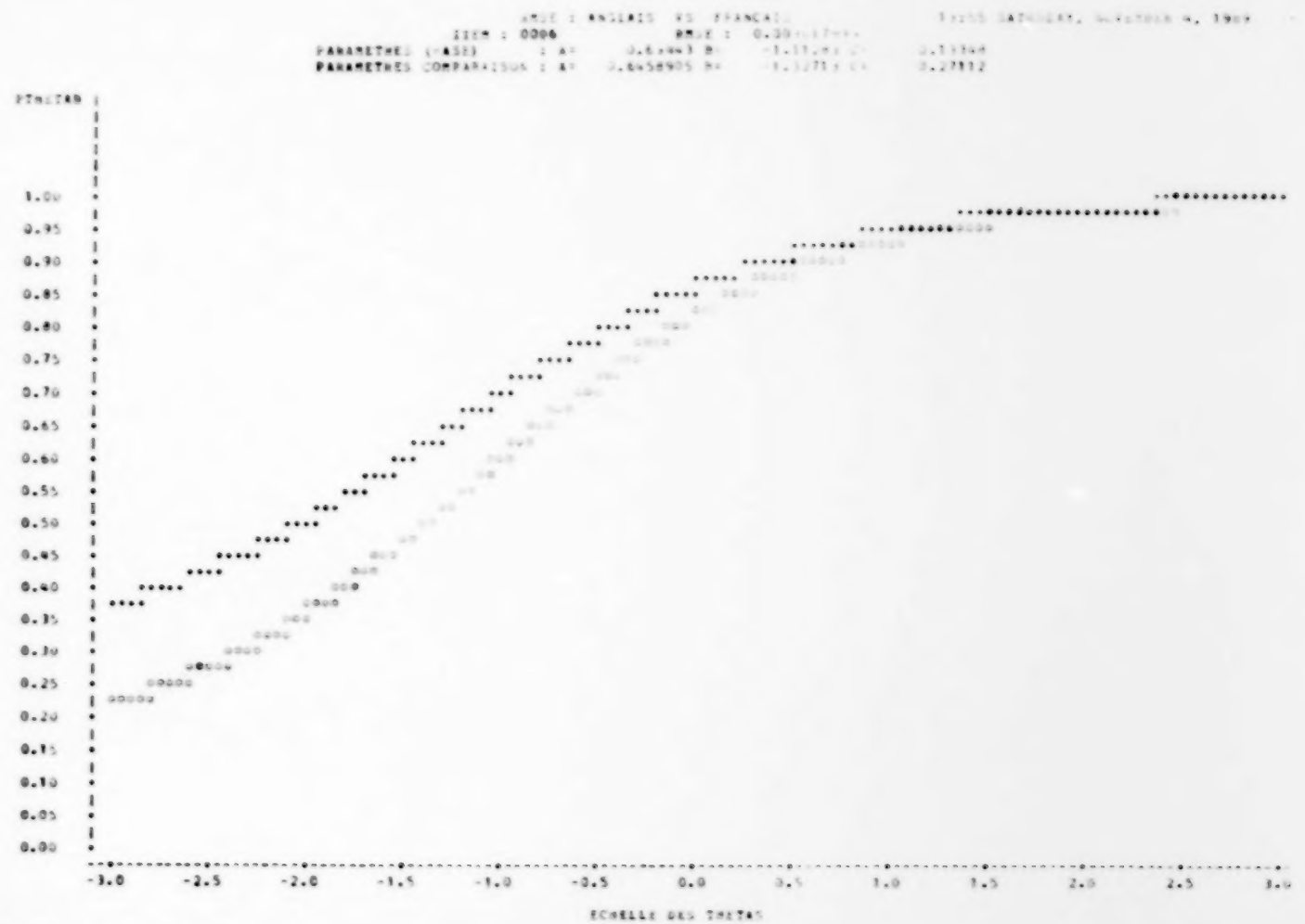
NOTE: 923 OBS HIDDEN

* version française
\+ version Anglaise

Annexe 4E


ANGLAIS VS FRANCAIS
ITEM : 0005
PARAMETRES (BASE) : A= 0.96228 B= 0.43236 C= 0.17579
PARAMETRES COMPARAISON : A= 0.8765073 B= 3.0844469 C= 0.22595
PTHETAS
1.00
0.95
0.90
0.85
0.80
0.75
0.70
0.65
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
-3.0
-2.5
-2.0
-1.5
-1.0
-0.5
0.0
1.0
1.5
2.0
2.5
ECHELLE DES THETAS
NOTE: 909 ORS HIDDEN
* version française
+ version anglaise

Annexe 4F



NOTE: 917 085 MIDDLE

* version française
\+ version anglaise

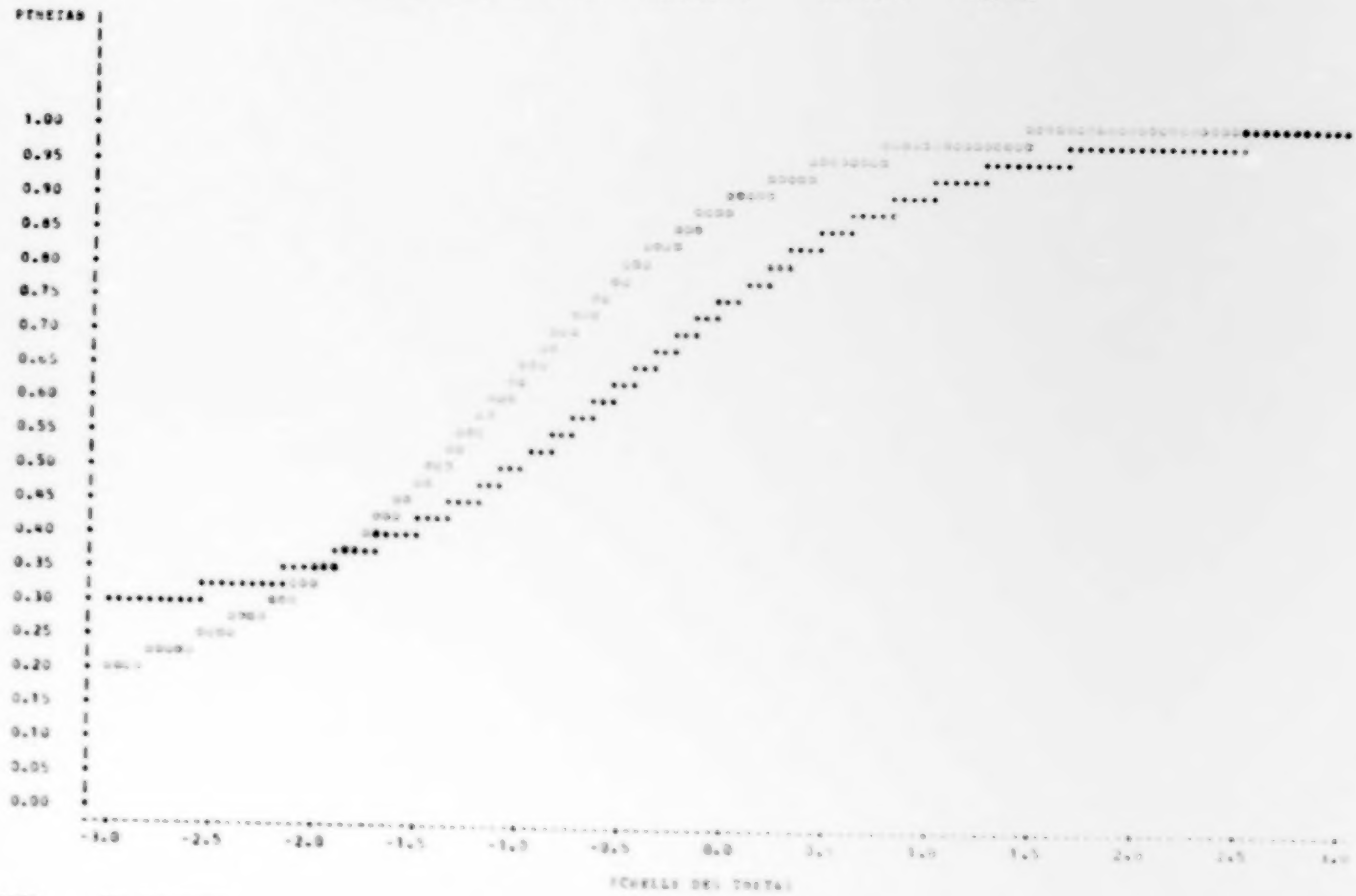
Annexe 4G

RMSE : ANGLAIS VS FRANCAIS

ITEM : 0037 RMSE : 0.00[illegible]

PARAMETRES (BASE) : A= 0.93808 B= -1.15742 C= 0.15631

PARAMETRES COMPARAISON : A= 0.8086321 B= -0.4516 C= 0.27452



* version française
+ version anglaise

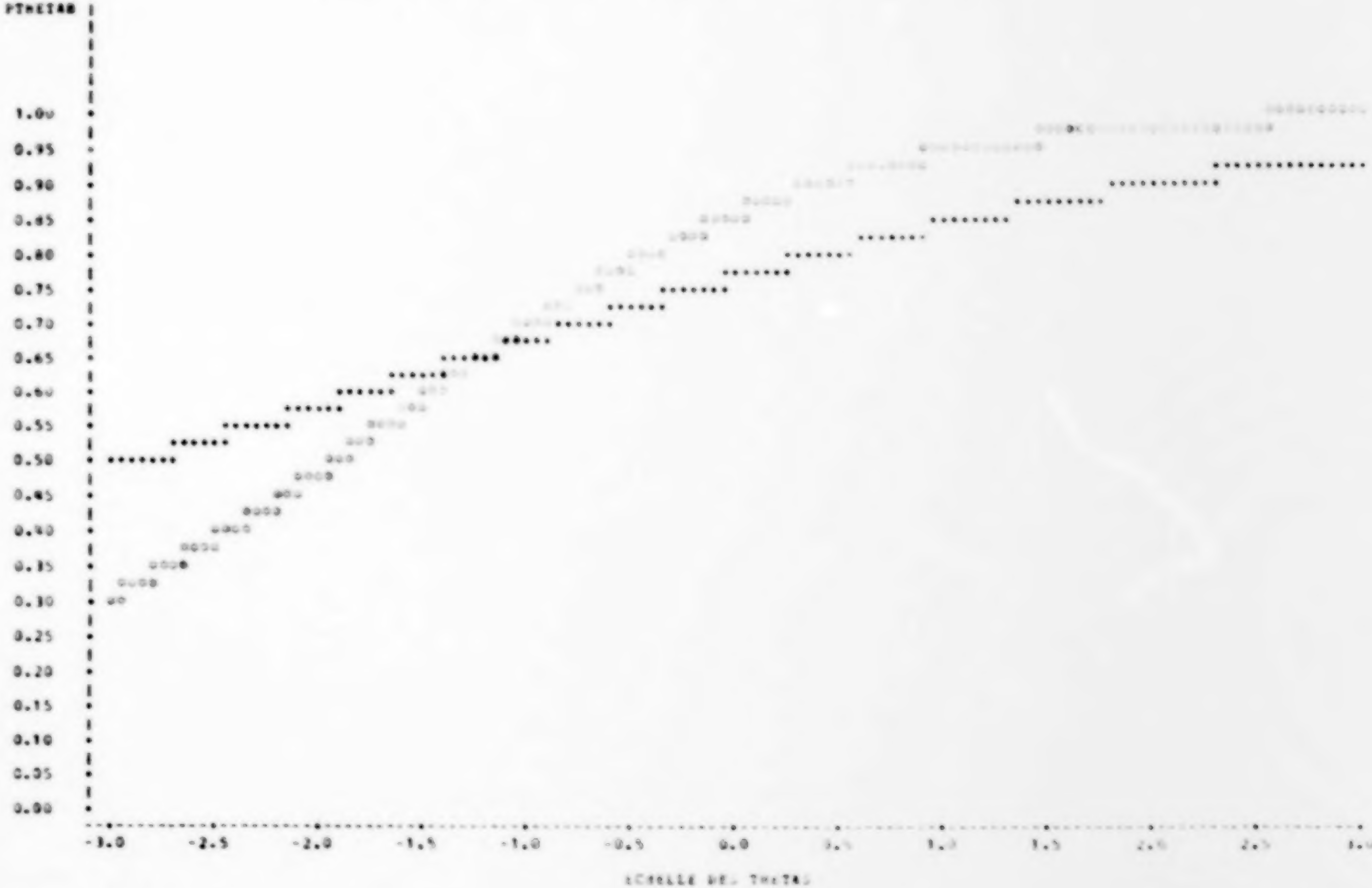
Annexe 4H

[illegible] : ANGLAIS VS FRANCAIS [illegible] SATURDAY, [illegible] 1989

ITEM : 0008 [illegible] : 0.00[illegible]

PARAMETRES (BASE) : A= 0.60195 B= -1.5[illegible] C= [illegible]

PARAMETRES COMPARAISON : A= [illegible] B= -1.[illegible] C= 0.26759



NOTE: 925 OBS HIDDEN

* version française
\+ version anglaise

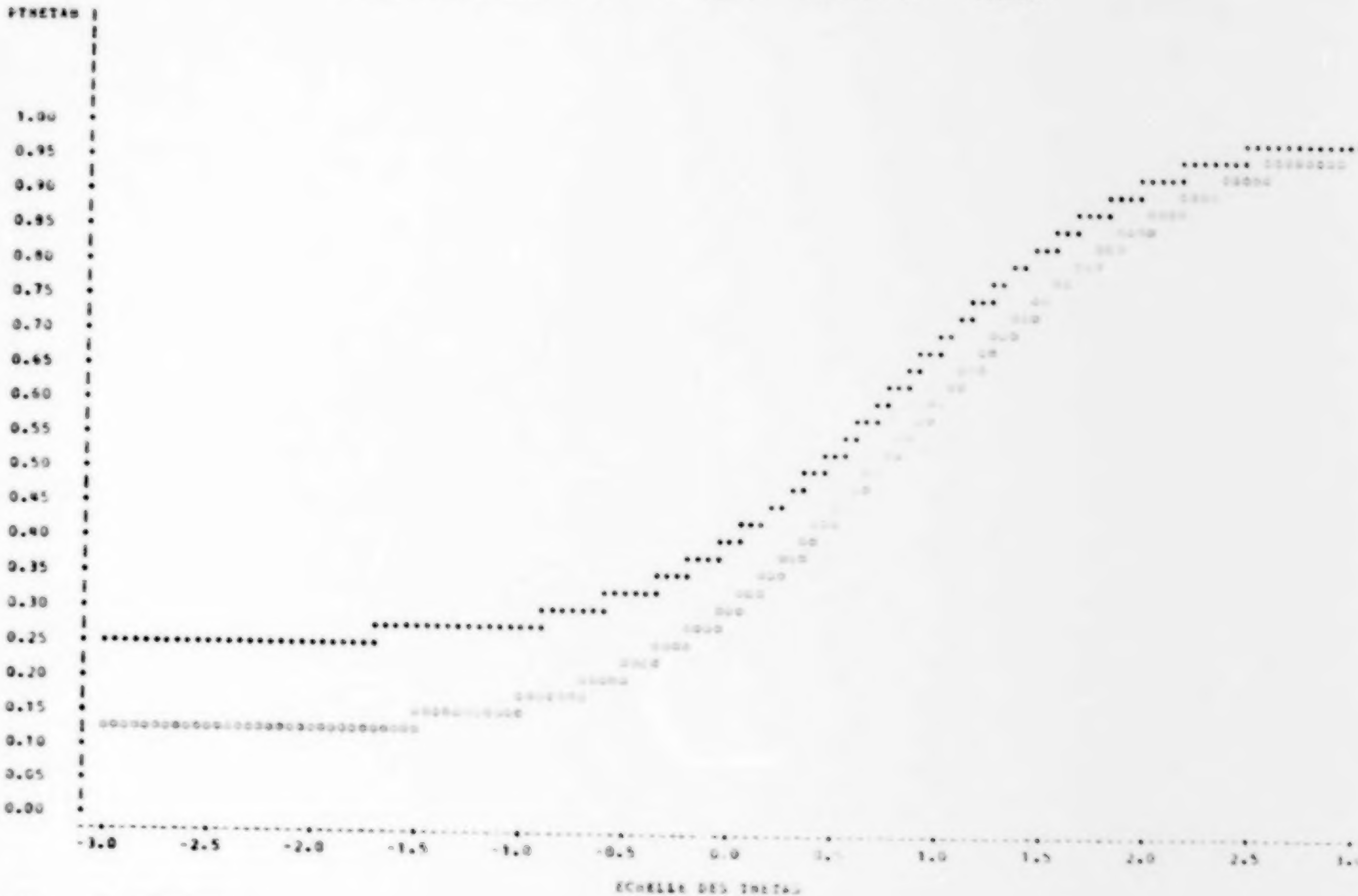
Annexe 41

RMSE : ANGLAIS VS FRANCAIS — 13:55 SATURDAY, NOVEMBER 4, 1989

ITEM : 0009 — RMSE : 0.0041[illegible]

PARAMETRES (BASE) : A= 0.98923 B= 0.86017 C= 0.11251

PARAMETRES COMPARAISON : A= 1.041252 B= [illegible] C= 0.25369

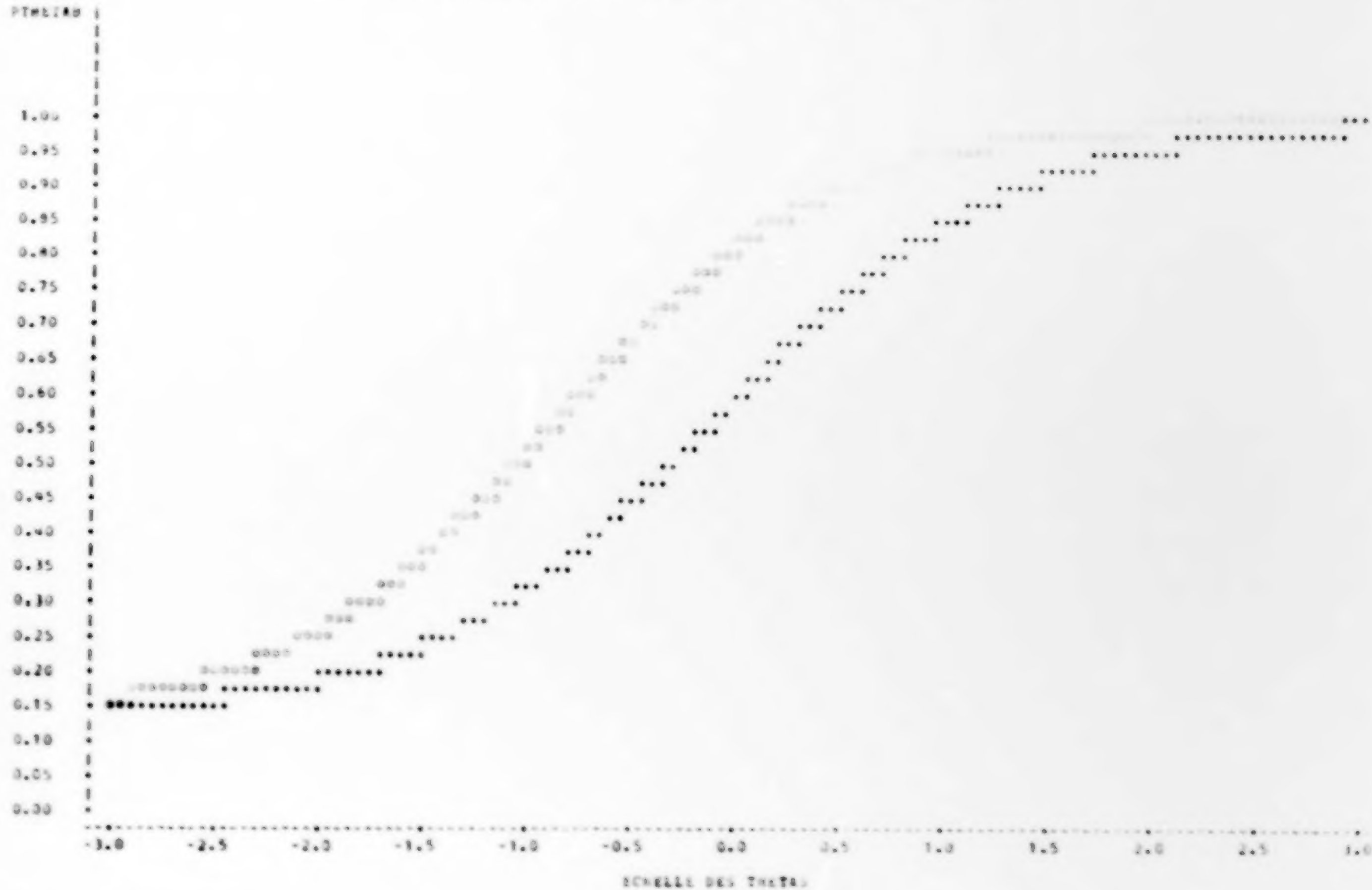


NOTE: 910 OBS HIDDEN

* version française
\+ version anglaise

Annexe 4J

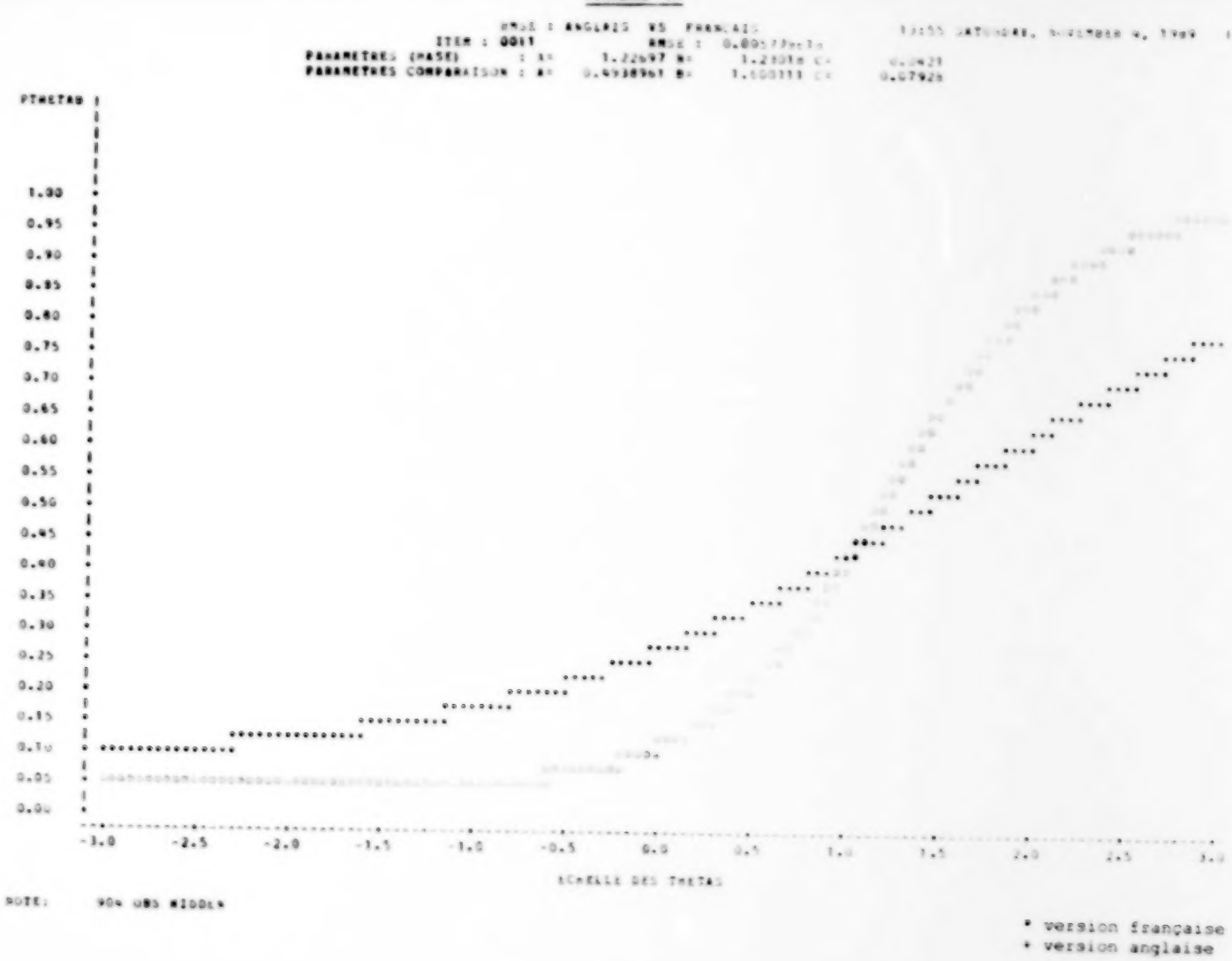
BASE : ANGLAIS VS FRANCAIS [illegible]
ITEM : 0010 RMSE : [illegible]
PARAMETRES (BASE) : A= 0.89177 B= -0.86318 C= 0.1243
PARAMETRES COMPARAISON : A= 0.8329413 B= -0.100558 C= 0.13161



NOTE: 90 OBS HIDDEN

* version française
* version anglaise

Annexe 4K



NOTE: 904 UBS MIDDLE

* version française
* version anglaise

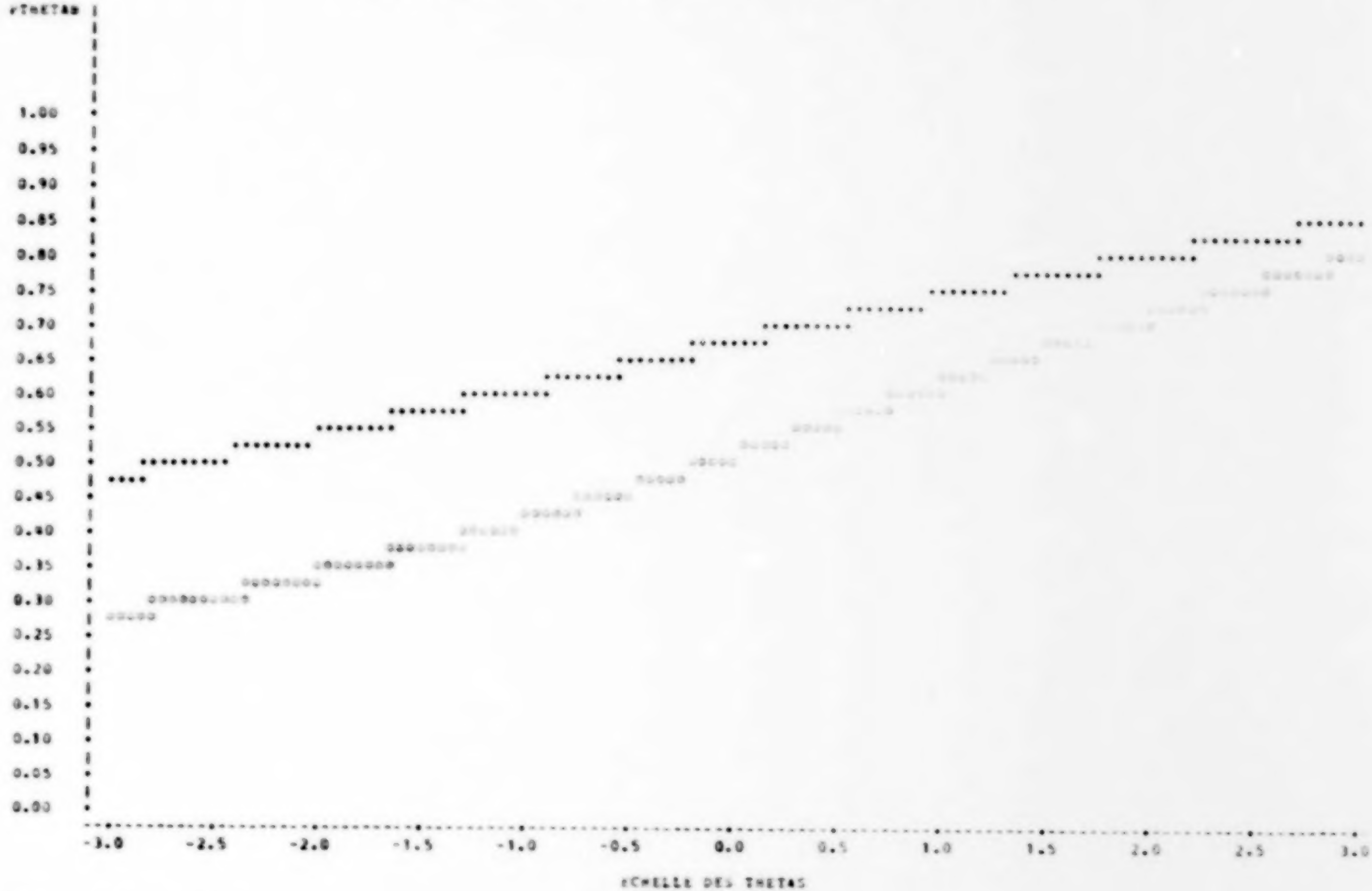
Annexe 4 L

RMSE : ANGLAIS VS FRANCAIS 13:55 SATURDAY, NOVEMBER 4, 1989

ITEM : 0012 RMSE : 0.00[illegible]

PARAMETRES (BASE) : A= 0.24613 B= 0.70052 C= 0.16655

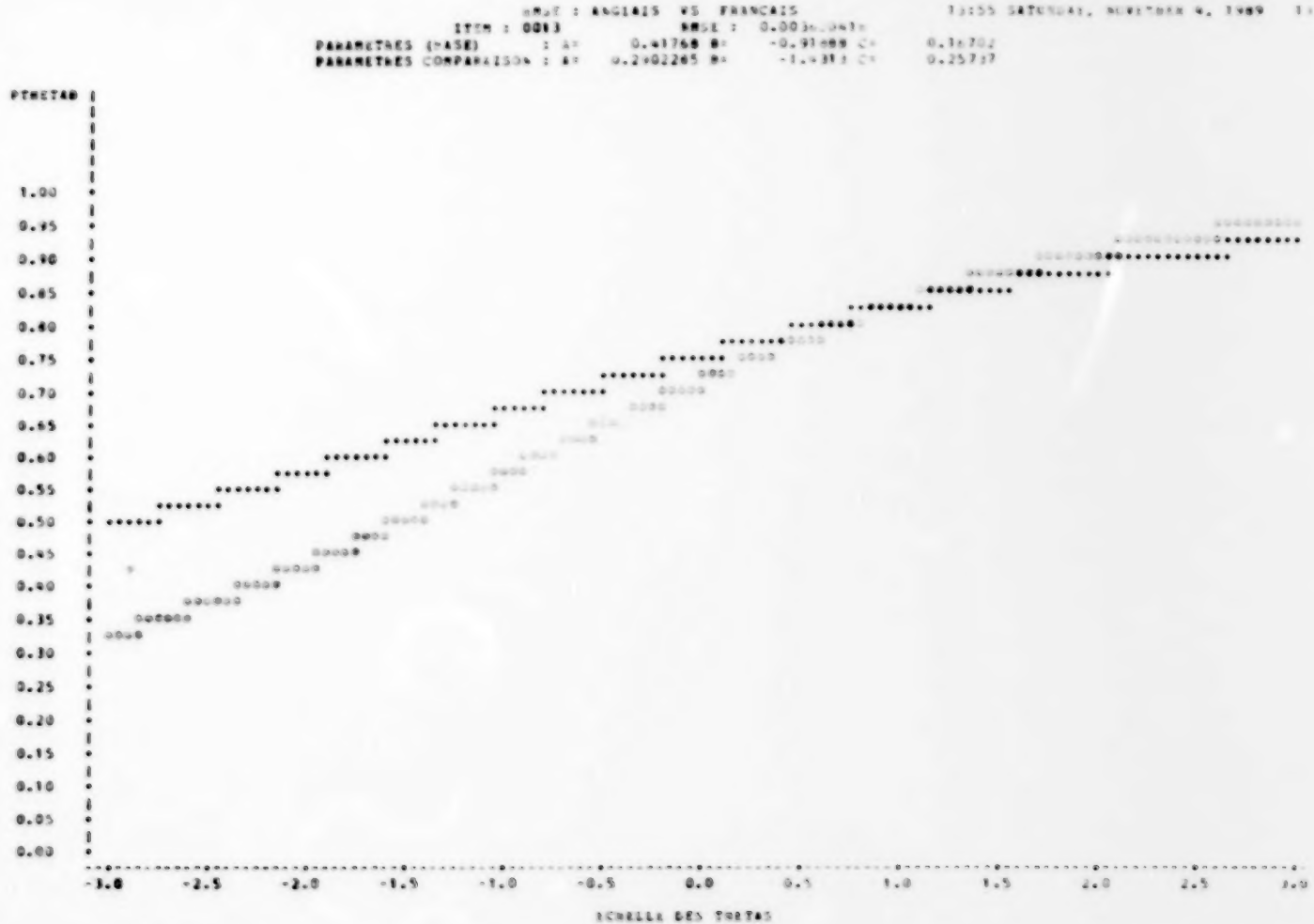
PARAMETRES COMPARAISON : A= 0.2219[illegible] B= -0.[illegible] C= 0.26335



NOTE: 936 OBS HIDDEN

* version française
\+ version anglaise

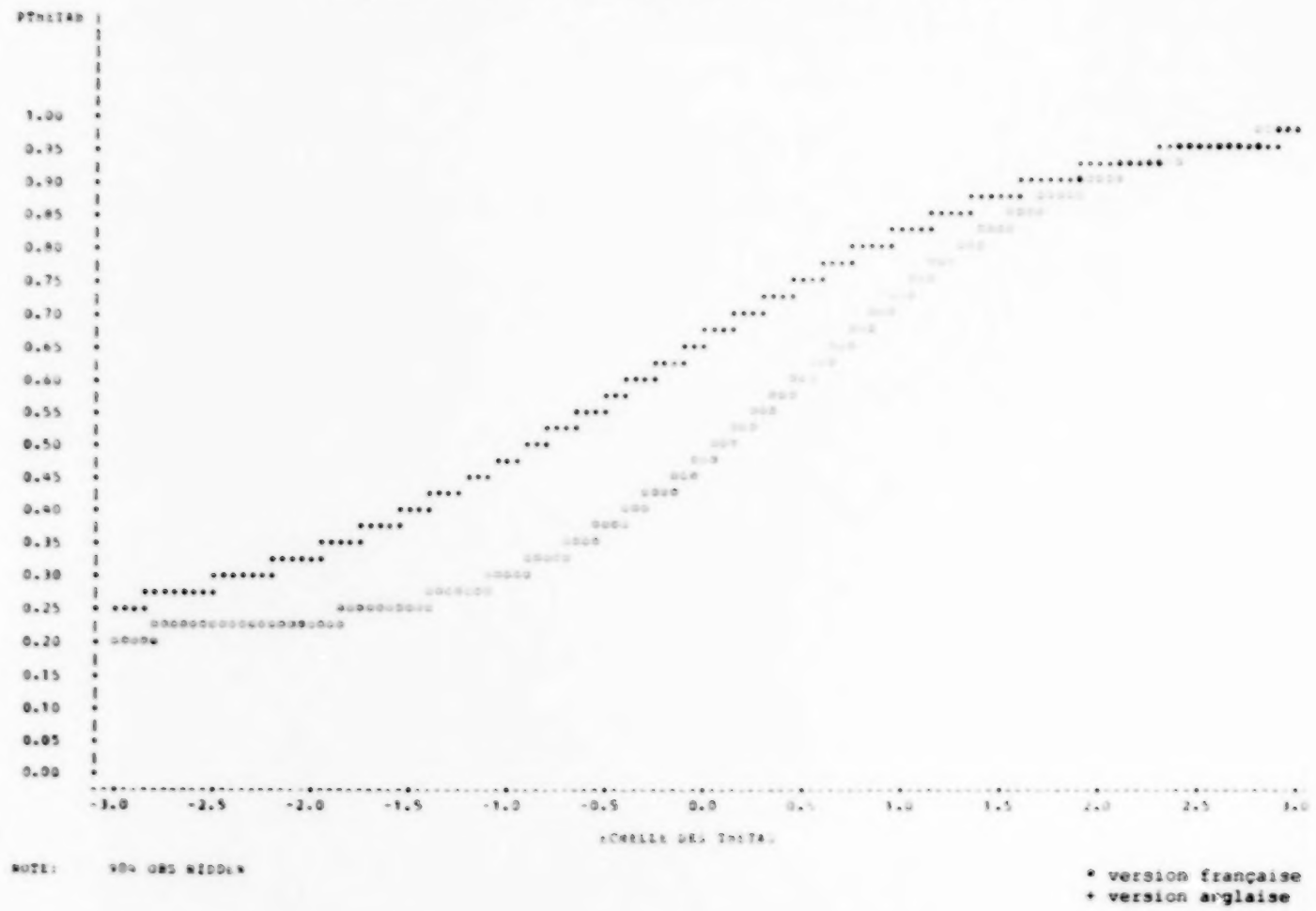
Annexe 4R



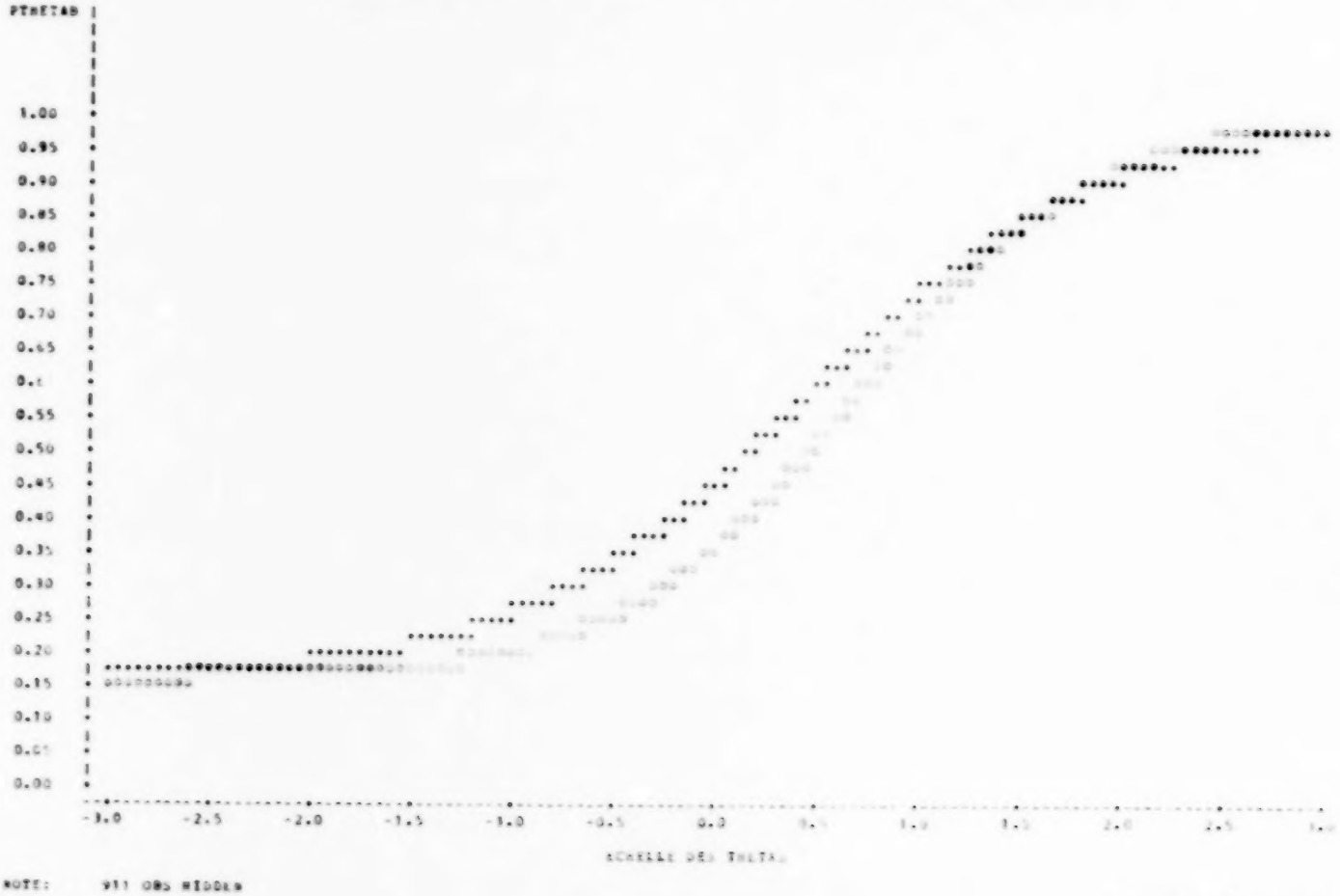
NOTE: 926 OBS RIDGE

* version française
\+ version anglaise

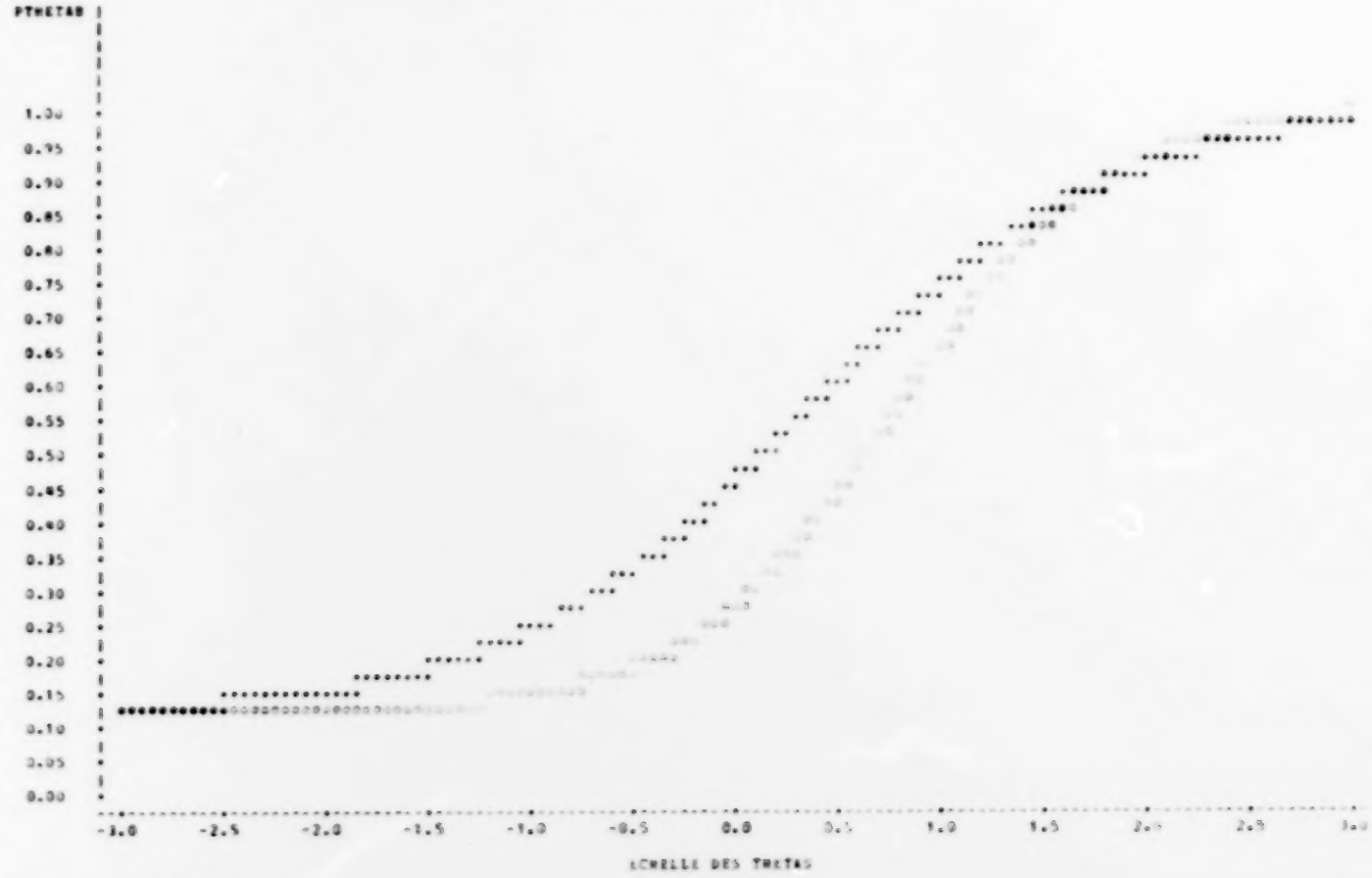
Annexe 4N


ITEM : 0014
PARAMETRES (BASE) : A= 0.77087 B= 0.59291 C= 0.20225
PARAMETRES COMPARAISON : A= 0.5509038 B= -0.337016 C= 0.19313
PTnbIAD
1.00
0.95
0.90
0.85
0.80
0.75
0.70
0.65
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
-3.0
-2.5
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
2.5
3.0
NOTE: 984 OBS MIDDLE
* version française
+ version anglaise

Annexe 40


TITRE : ANGLAIS VS FRANCAIS
ITEM : 0015
PARAMETRES (BASE) : A= 1.01441 B= 0.69092 C= 0.15972
PARAMETRES COMPARAISON : A= 0.8140208 B= 0.4504956 C= 0.16303
PTHETAB
1.00
0.95
0.90
0.85
0.80
0.75
0.70
0.65
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
-3.0
-2.5
-2.0
-1.5
-1.0
-0.5
0.0
ECHELLE DES THETA
NOTE:
* version française
+ version anglaise

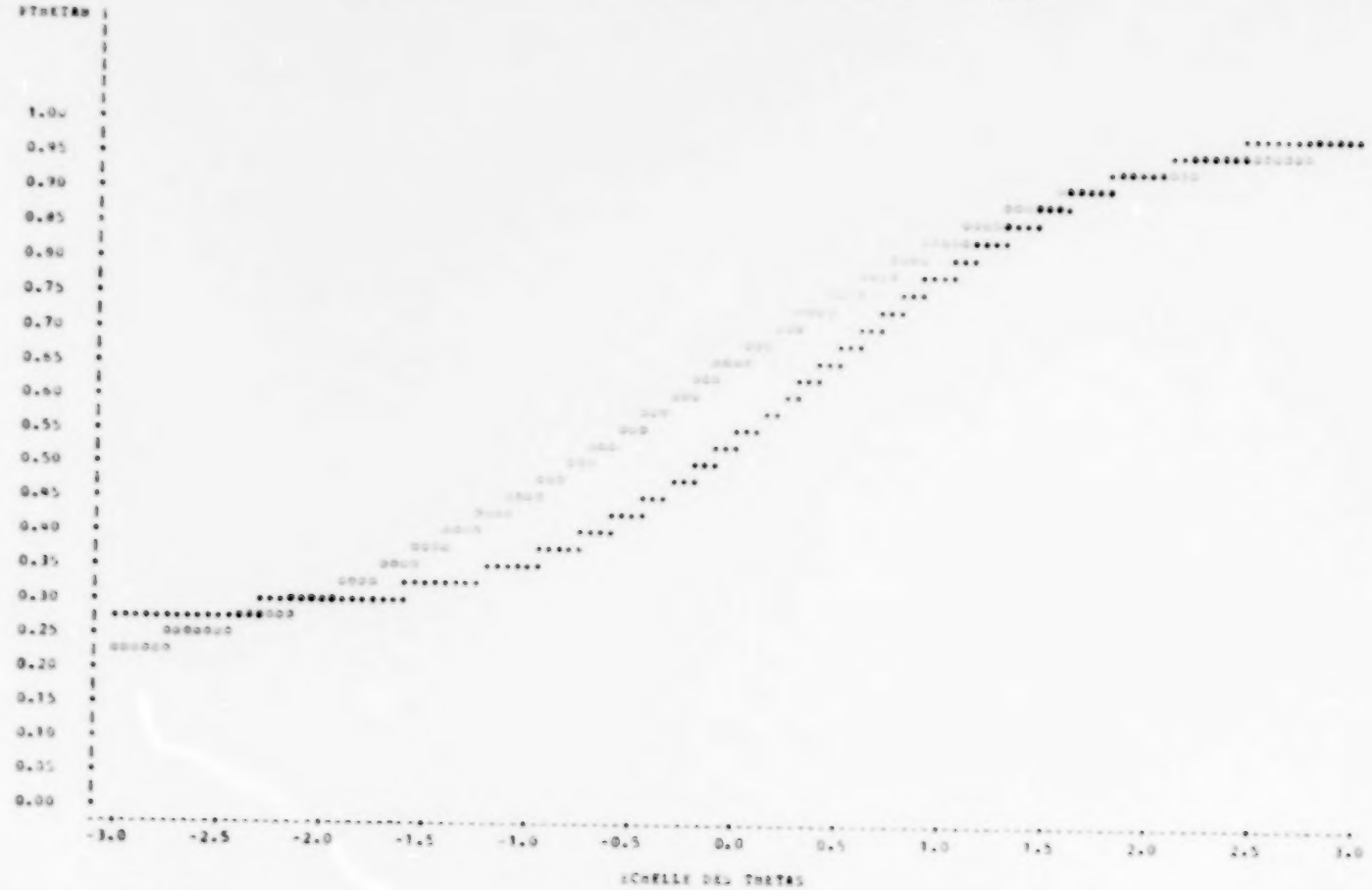
Annexe 4P



NOTE: 900 OBS MISSING

* version française
\+ version anglaise

Annexe 40



NOTE: 915 OBS MISSING

* version française
\+ version anglaise

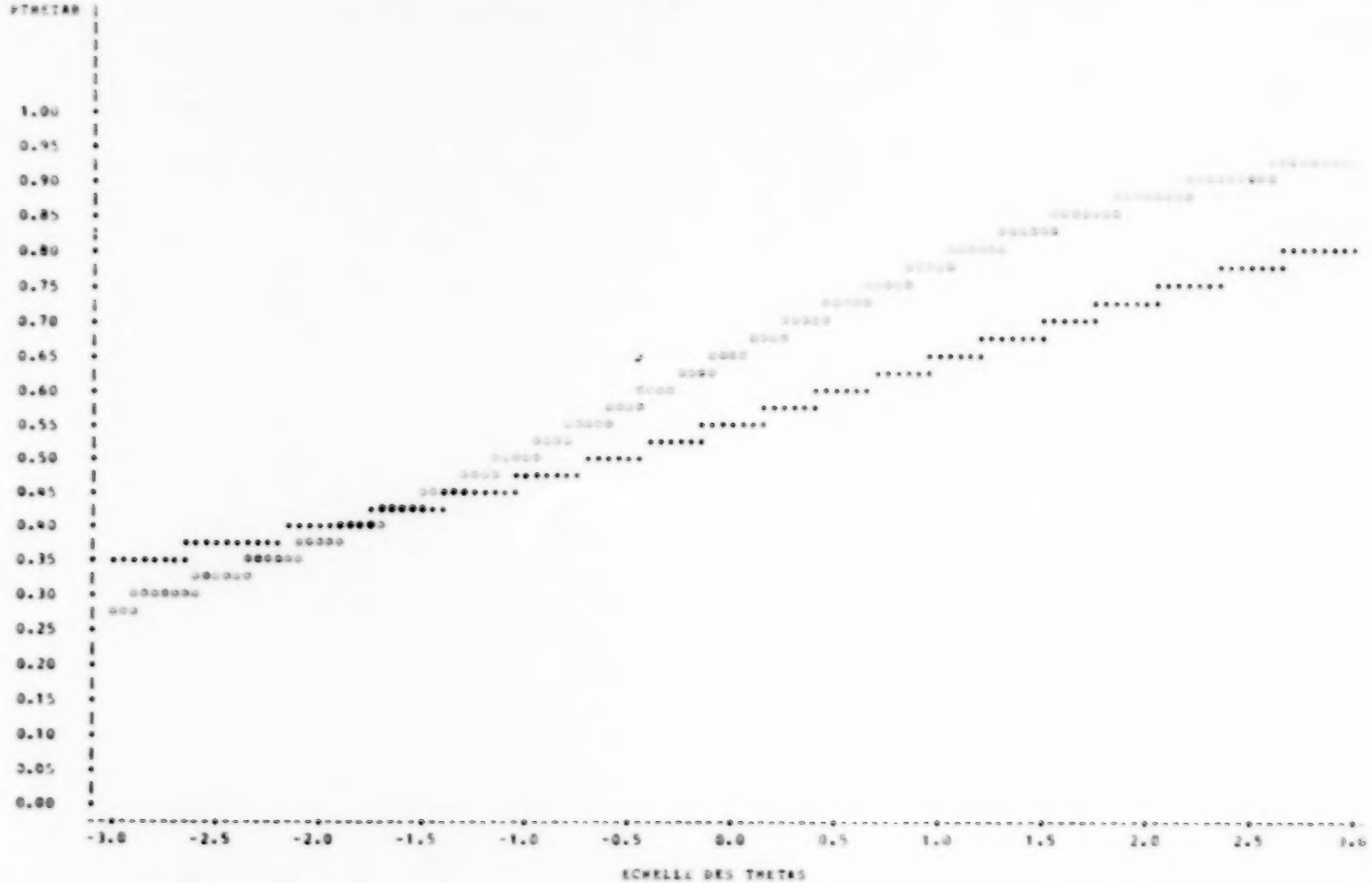
Annexe 4R

RMSE : ANGLAIS VS FRANCAIS 11:55 SATURDAY, NOVEMBER 6, 1999 11

ITEM : 0018 RMSE : 0.0061.2237

PARAMETRES (BASE) : A= 0.80641 B= -0.53420 C= 0.19019

PARAMETRES COMPARAISON : A= 3.2935724 B= 0.83709 C= 0.24936



NOTE: 922 OBS MIDDLE

* version française
\+ version anglaise

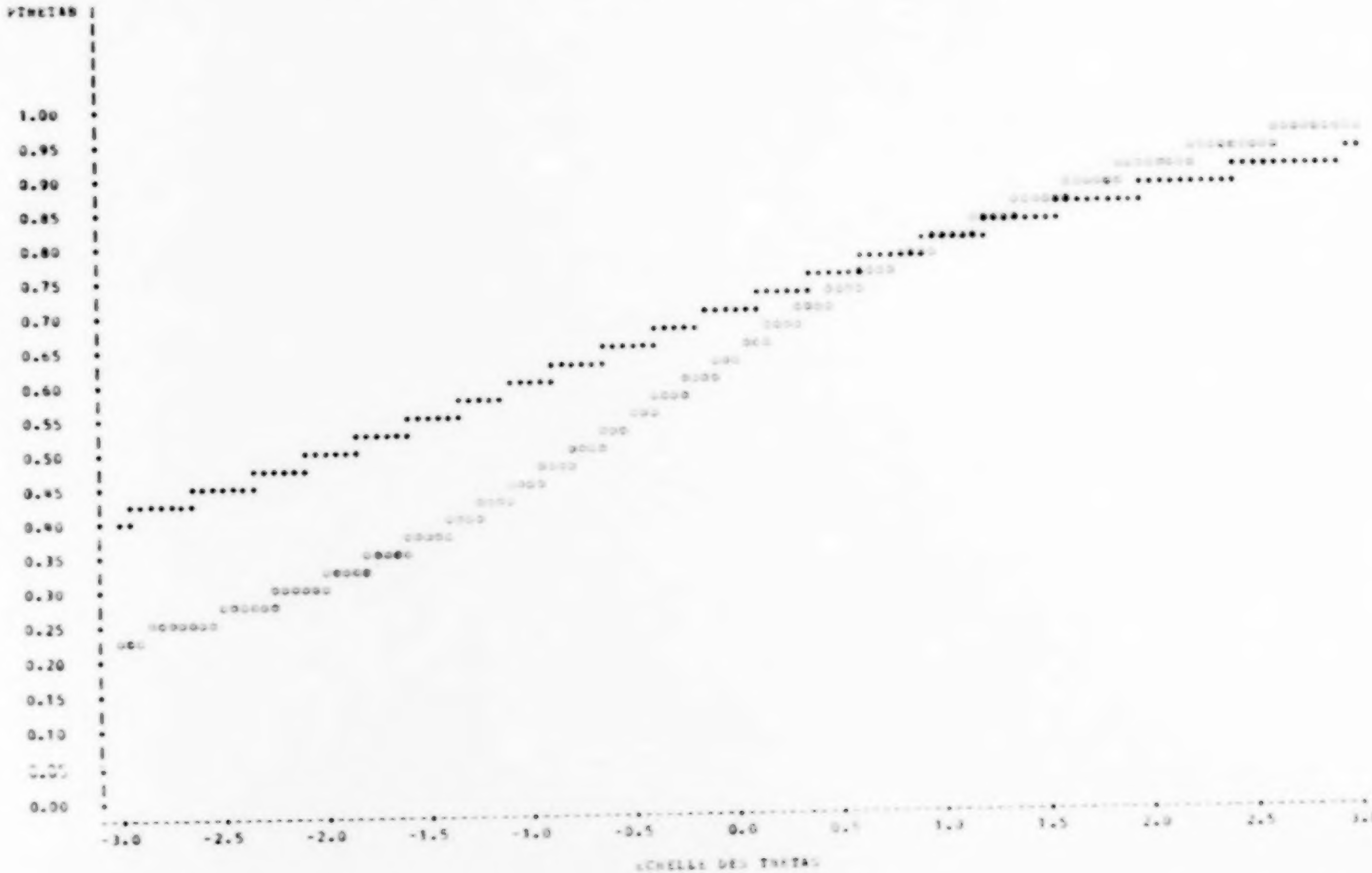
Annexe 45

RMSE : ANGLAIS VS FRANCAIS   13:55 SATURDAY, NOVEMBER 4, 1989 1

ITEM : 0019   RMSE : 0.0046[illegible]

PARAMETRES (BASE) : A= 0.51186 B= -0.23771 C= 0.15649

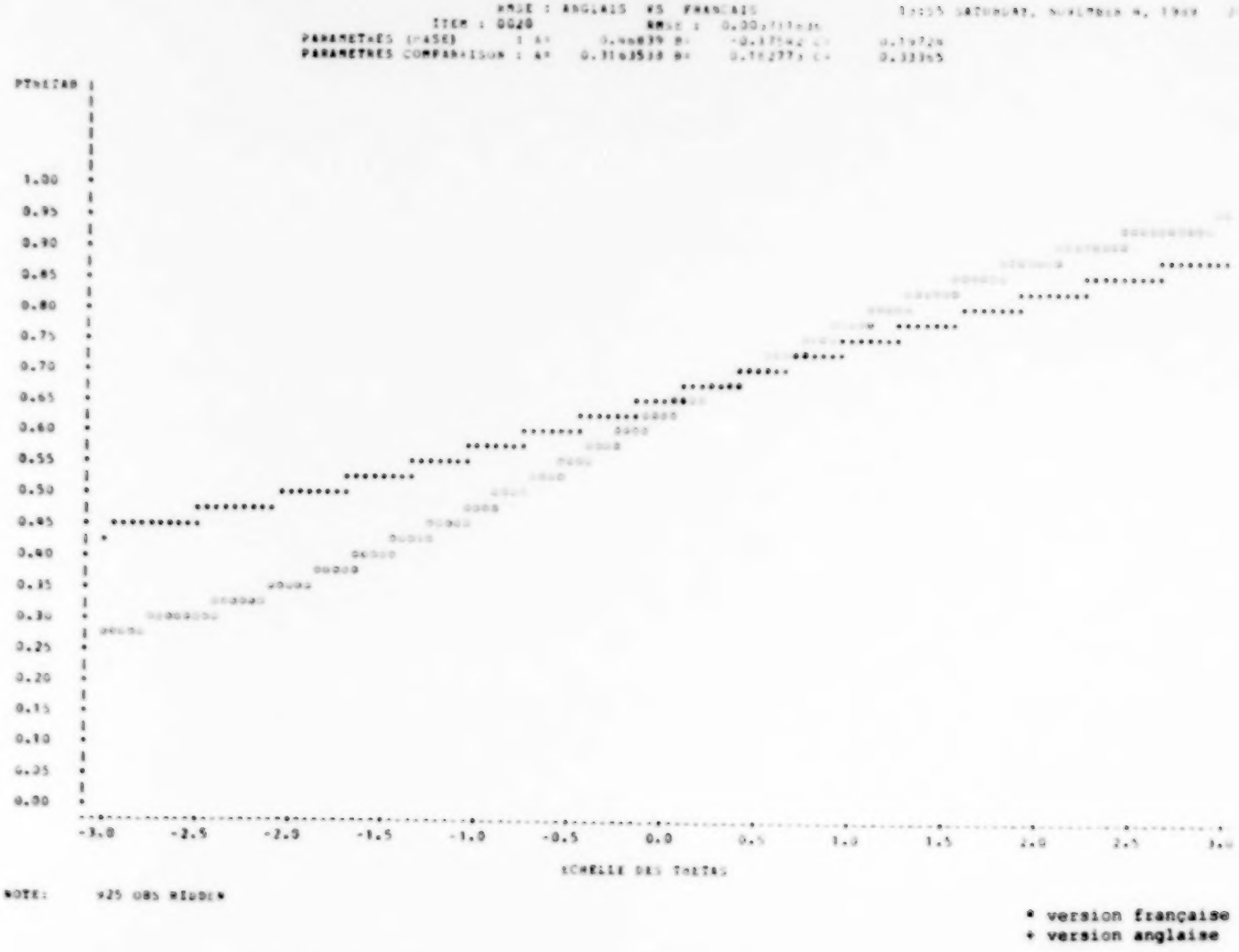
PARAMETRES COMPARAISON : A= 0.3157464 B= -0.460005 C= 0.20997



NOTE: 918 OBS HIDDEN

* version française
\+ version anglaise

Annexe 4T


ANGE : ANGLAIS VS FRANCAIS
ITEM : 0020
PARAMETRES (=ASC)
PARAMETRES COMPARAISON
PTBISAB
1.00
0.95
0.90
0.85
0.80
0.75
0.70
0.65
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
-3.0
-2.5
-2.0
-1.5
-1.0
-0.5
0.0
ECHELLE DES TOTAL
NOTE:
* version française
+ version anglaise